Indoor-Outdoor Detection and Routing Using Bayesian Networks

Juan José Balcázar

School of Electrical Engineering

Espoo 20.6.2016

Thesis supervisor:

Prof. Simo Särkkä

Thesis advisor:

M.Sc. (Tech) Ville Tolvanen

Aalto University
School of Electrical Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

Author: Juan José Balcázar

Title: Indoor-Outdoor Detection and Routing Using Bayesian Networks

Date: 20.6.2016 Language: English Number of pages: 16+81

Department of Electrical Engineering and Automation

Professorship: Sensor informatics and medical technology

Supervisor: Prof. Simo Särkkä

Advisor: M.Sc. (Tech) Ville Tolvanen

The purpose of this Master's thesis is to do research on methods for indoor-outdoor detection and to implement an Android application for the navigation through indoor and outdoor environments. Due to the length and complexity of the project, it has been divided in two different parts: the outdoor navigation with the indoor-outdoor detection and the indoor guidance. Here we are going to explain the first part, beginning with the study of recent papers about outdoor positioning and detection techniques. On the one hand the popular method to locate yourself outdoors is the GPS that has proved to be accurate and robust for path planning. On the other hand current studies of indoor-outdoor detection are based on the light intensity, the magnetic field, the cell tower strength, the number of satellites and the signal to noise ratio of the satellites. We have used those parameters to build a Bayesian network that uses their conditional probabilities to get the probability of being outdoors. Then we have implemented the outdoor navigation and the detection algorithm with the indoor guidance in the Android application. The indoor positioning system was provided by IndoorAtlas that uses the geomagnetic disturbances of buildings to locate the user indoors. To do that, the system requires a geomagnetic map of the building that has to be made beforehand. For this reason the application has been built only for a certain shopping center that was already mapped. With all this, the final Android application is able to guide the user from a store inside the mall to a place outdoors and vice versa using a combined indoor and outdoor guidance.

Keywords: Indoors, outdoors, detection, navigation, Android.

Acknowledgements

I would first like to express my deep gratitude to my thesis supervisor Prof. Simo Särkkä. He allowed this paper to be my own work, but helped me out whenever I needed advice. I would also like to thank my thesis advisor M.Sc. (Tech) Ville Tolvanen for guiding me throughout the project to reach the initial objectives.

This thesis would not have been possible without the resources provided by the company IndoorAtlas that has supported the project from the beginning. I would also like to acknowledge Aalto University and my home university Universidad Politécnica de Madrid for giving me the opportunity of doing my master thesis in a foreign country.

Special thanks are given to my two flatmates, partners and friends Jose Miguel Indave and Daniel Herrero. Particularly to Dani for being the other half of the project always encouraging me to research and write this thesis.

Finally, I must express my gratitude to my parents and to my girlfriend for providing me with unfailing support and continuous inspiration from the distance.

Otaniemi, 20.6.2016

Juanjo B. S. Engineer

Contents

List of Figures

Symbols and abbreviations

Abbreviations

API	Application Programming Interface
CDMA	Code Division Multiple Access
GPS	Global Positioning System
GSM	Global System for Mobile communications
IPS	Indoor Positioning System
NFC	Near Field Communication
NMEA	National Marine Electronics Association
RSS	Received cellular Signal Strength
WLAN	Wireless Local Area Networks

1 Introduction

The aim of this project is to study methods for indoor-outdoor detection and to implement an Android application capable of planning a route through indoor and outdoor environments. It is a very large project so the work has been divided in two distinct parts: 1. the outdoor positioning and routing with the indoor-outdoor transition and 2. the indoor positioning and routing. In this Master's thesis we are going to explain the first part of indoor-outdoor transition and outdoor guidance. Because the purpose of the project is to build a single application, we will also explain how this part is combined with the indoor guidance part (developed in other Master's thesis).

First of all we are going to take a look to the current methods of outdoor and indoor positioning. The most common way to locate yourself in outdoor environments is the GPS that uses the satellites placed around the Earth and a receiver to get an accurate location of that receiver on the Earth's surface (Zogg, 2002).

Regarding the indoor positioning and routing, we cannot use the GPS inside buildings because the signal of the satellites is weak. Recent studies use Indoor Positioning Systems (IPS) based on different technologies to achieve an indoor location (Arfwedson and Berglund, 2015). The most important techniques are those which use WiFi networks, radio waves, bluetooth beacons, and geo-magnetic fingerprinting (Chung et al., 2011). We are going to use the geo-magnetic fingerprinting supported by the company IndoorAtlas that has provided us the resources needed for developing the indoor navigation for the final Android application.

The indoor-outdoor detection is the main part of the thesis so we are going to look into the previous studies that have found a reliable solution. As we are developing a mobile phone application, the literature study focuses on papers that uses common cell phone sensors such as the light sensor, the magnetometer or the GPS. The starting point of our research was the preliminary research work done in IODetector (Zhou et al., 2012) that mainly uses the light sensor, the magnetometer and cell tower signals. Then, we did some experiments to gather information about the sensors that are relevant for indoor-outdoor detection. The result of the experiments showed that the critical parameters are the light intensity, the magnetic field difference, the number of satellites and the signal to noise ratios (SNR) of the satellites.

With the above mentioned parameters we developed an algorithm based on Bayesian networks that has the probability of being outdoors as outcome. We selected the Bayesian network system because it provides a graphical solution based on the conditional probabilities of the components involved (Charniak, 1991). These conditional probabilities are continuous and depend directly on the values measured by the smartphone. The probability of being outdoors is a value between 0 and 1, the higher the value the more certain the algorithm is of being outdoors. We used a threshold of 0.5, meaning that higher values consider that the device is outdoors.

Once we found a reliable algorithm for the indoor-outdoor detection, we moved on the implementation of the outdoor guidance and indoor-outdoor transition on our Android application. The outdoor navigation uses Google Maps API that contains all the resources needed. Using this API we can place the current location of the mobile phone on a map and draw a route from that position to any desired location. For the indoor-outdoor transition we required a really robust detection and therefore we have evaluated all the possible scenarios that may affect to the algorithm: GPS enabled during the day, GPS enabled during the night, GPS not enabled during the day and GPS not enabled during the night.

We realized that when the GPS is not enabled during the night, the algorithm is weak because it only depends on the magnetic field difference that it is not a really accurate parameter. Thus, we added the cell strength as an extra parameter to the Bayesian network that can correct the probability of being out.

At this point all different parts of the application are implemented: the outdoor navigation, the indoor-outdoor transition and the indoor navigation. However the IndoorAtlas system requires a geomagnetic map of the building to get the real time indoor location. Therefore we only implemented the indoor guidance in the Sello mall.

The final application has two different interfaces depending on whether you are inside the shopping center or not. If you are inside Sello, you can either go to a store in Sello using the straight indoor navigation or go somewhere else using the indoor-outdoor guidance as follows. First you have to write the place where you want to go, then the indoor navigations leads you to the exit of the mall. When the application detects you are outdoors, it asks you to switch to the outdoor guidance that calculates the outdoor route from your current location to the desired goal using the Google API.

On the other hand, if you are outside Sello you can either go to a store in Sello using the outdoor-indoor guidance or go somewhere else using the Google Maps application. For the outdoor-indoor guidance the first step is to select the store inside Sello you want to go. Afterwards the outdoor navigation leads you from your current location to the entrance of the mall. As before, when the application detects you are inside Sello, it asks you to switch to the indoor guidance that calculates the indoor route from your current location to the shop selected.

2 Background

The purpose of the final application is to achieve an indoor-outdoor guidance. The first step is to detect somehow if you are indoors or outdoors and then carry out the guidance in both indoor and outdoor environments. To accomplish our goal, we are going to study similar projects and different technologies that can help us to develop our detection system and take a look into the outdoor positioning and routing methods afterwards.

2.1 Outdoor Positioning and Routing

Mankind has always had the need to be located in the environment around. This need began years ago with cartographic maps that today have been overshadowed by new technologies such as computers and smartphones.

Nowadays, you can carry all the maps of the world in your pocket and locate yourself on them using the Global Positioning System or GPS, available in any current mobile phone. The GPS uses the satellites placed around the Earth to establish a position on the globe (Figure 1), explained in Section 2.1.1.

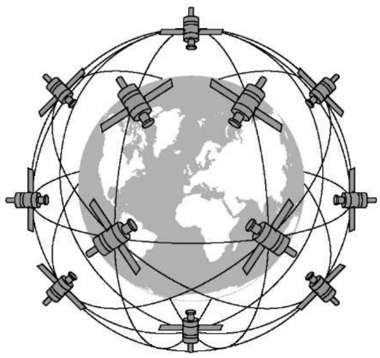

Figure 1: There are several satellites placed around the Earth. (Source: http://what-when-how.com)

This system works really well for outdoor positioning because the GPS receiver can see the satellites without obstacles, getting most of the signal emitted. However in indoor environments the signal is weaker than outdoors due to the presence of walls and other structures that may disturb or interrupt the signal (Zogg, 2002).

Apart from the GPS, you can locate yourself in outdoor environments using WiFi, a technology that allows electronic devices to connect to a Wireless Local Area Network (WLAN). The following sections describes both of the systems from a technical point of view.

2.1.1 GPS

The GPS technology is the most common way to locate yourself in outdoor environments. This system uses the satellites in the Earth's orbit to measure how far are they from the receiver device and get the location on the surface.

This works as in Figure 2: the satellites around the planet are constantly emitting their clock time and position and then a GPS receiver measures the signal transit between the point of observation and four different satellites to get the current longitude, latitude, altitude and time of the receiver (Zogg, 2002).

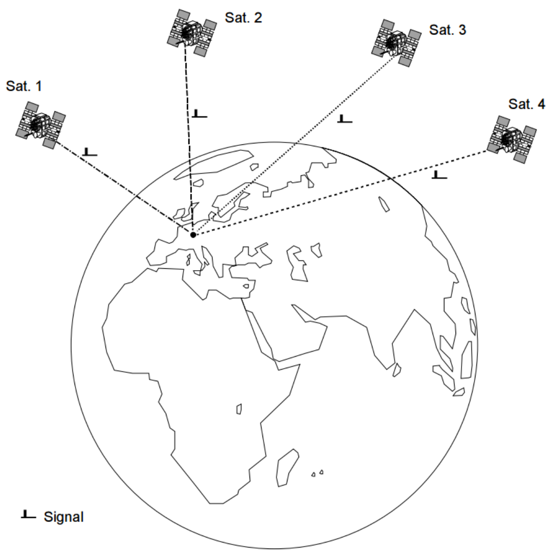

Figure 2: The position of the receiver is determined with the information provided by the satellites (Zogg, 2002).

However, the functionality of this positioning system is limited due to the communication between the device and the satellites used around the planet. This means the positioning is accurate enough only when the device can see the satellites. Thus, the GPS is a good system for outdoor positioning and planning but not for indoor positioning because the receiver may lose the signal of the satellites due to the walls and ceiling of buildings.

This technology is highly developed and there are several applications that use it. Google Maps for example is a web mapping service developed by Google which not only provides a location in the world using GPS but also offers several features such as satellite imagery, street maps, real-time traffic conditions and path planning (by foot, car, public transport or even bicycle).

Google Maps has become the most popular location service of the world thanks to the free and open-source policy of the company that allows developers to use the

Google Maps Application Programming Interface or API to implement Google Maps into their websites or applications.

For our smartphone application, we will use this Google Maps Direction API because simplifies the code and is well established.

2.1.2 WiFi

Even though GPS is nowadays the most popular form of positioning system, WiFi positioning has received much attention in recent years. This is due to the presence of more and more WiFi access points and its duality to perform well in outdoor and indoor environments (Li et al., 2008). Current research focuses on two different approaches to get the location based on the WiFi network: trilateration and fingerprinting.

2.1.2.1 Trilateration

This technique is intuitive and is similar to how the GPS works. Starting from several WiFi access points where we know exactly their position in a certain area, a WiFi receiver can measure the distance to each of the access points.

So for each access point $A_1, A_2, ..., A_n$, the receiver may be at a distance of R_i $(i = 1, ..., n)$ in any direction building a circle around every access point. Those circles intersect at one single point which is the receiver position (Li et al., 2008). Notice that the minimum number of access points in range needed are three as shown in the image below.

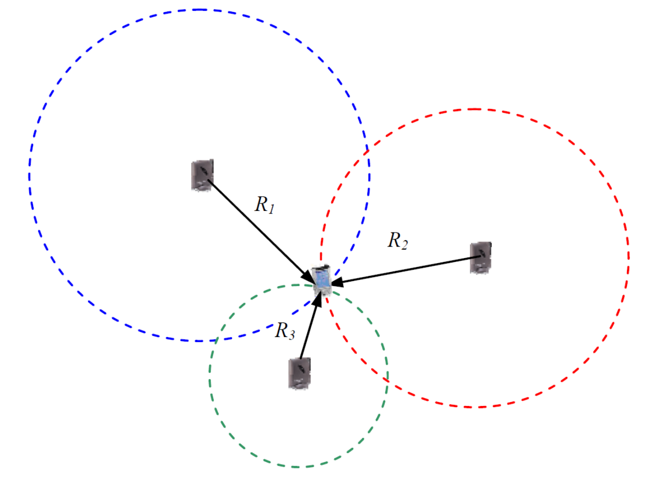

Figure 3: Trilateration approach using WiFi (Li et al., 2008).

2.1.2.2 Fingerprinting

The point of Wifi fingerprinting is to map the measured radio signals of the relevant locations for the positioning. There are two phases in the fingerprinting approach: offline and online phases (Mok and Retscher, 2007). The purpose of the offline phase is to build a database and train the model. To achieve that, a certain amount of locations will be carefully selected and then a number of calibration measurements will be performed. With these measurements a histogram is made and stored in the database. In the online phase the measurements made by a WiFi receiver (smartphone) are compared against the values of the database using an appropriate search/matching algorithm. When the measurement is found, the corresponding location of the database is given (Li et al., 2008).

2.2 Indoor Positioning and Routing

As we said in the previous paragraphs, the traditional GPS system is not accurate enough for the positioning and routing in indoor environments. For this reason the recent research focuses on new ways to locate yourself inside buildings. These methods are called indoor positioning systems (IPS) which are used to place people inside buildings where the GPS does not work properly (Arfwedson and Berglund, 2015).

There are different ways to develop the IPS using distinct kinds of technology, here we present the most important ones that can be applied to the smartphones.

- **WLAN / WiFi:**

  WLAN (Wireless Local Area Networks) can be used to estimate the location of a mobile device within this network. The use of this technology is very useful because WLAN access points are available in many indoor environments such as hospitals, universities and malls. For this reason, WLAN indoor positioning systems have become the most widespread approach for indoor localization (Mautz, 2012).

  However the accuracy of the system is not as good as other systems with a 20-50 meters range and needs a right infrastructure to support the connection between router and mobile device.

- **RFID:**

  RFID (Radio Frequency IDentification) technology is used to send and receive data via radio waves. An RFID system is basically a reader (RFID scanner) that receives the data sent from the RFID tags. This is used to get the position of an RFID tag based on the proximity information of the data transmitted, also known as CoO (Cell of Origin) (Mautz, 2012).

Nowadays most of the cellphones can use this technology for positioning in indoor environments. The accuracy is directly related to the density of tag deployment and reading ranges (Liu et al., 2007), performing good results only in certain places.

- **Beacons:**

  A beacon is a device that can communicate with other devices using infrared or bluetooth technology to share information Mautz (2012). To locate the smarphone on a building several beacons are needed placed at known locations inside the building. The idea is similar to the GPS system explained in Section 2.1. The smartphone gets the signal intensity of at least 3 beacons and calculates the distance to locate the smartphone on the known map of the building (Chawathe, 2008).

  The accuracy of this system depends on the number of beacons placed in the building, meaning that it needs a quite big invest on infrastructure to make the system accurate enough for the location purpose.

- **Geo-Magnetic Systems:**

  The principle of these systems of navigation inside buildings is to use the Earth's magnetic field which acts like a huge magnet.

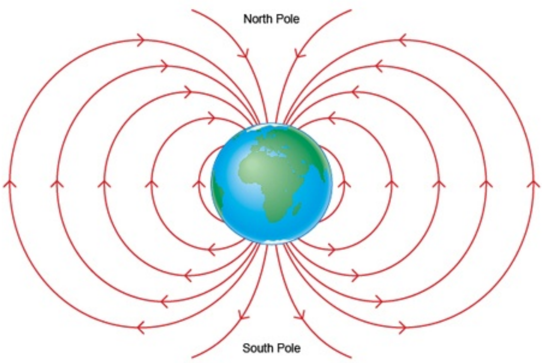

Figure 4: The Earth acts like a magnet (Mautz, 2012).

  Each location of the planet has a non-zero magnetic flux density. Inside buildings these values vary due to man-made sources such as metal building materials, electric power systems and industrial devices. The systems which uses this approach use the fluctuations of the magnetic field measured by a magnetometer to make a magnetic map of the building (Haverinen and Kemppainen, 2009).

  Unlike other technologies, geo-magnetic localization does not require any infrastructure and the accuracy relies on how well the magnetic map is made (Chung et al., 2011).

This project was developed thanks to the support of the company IndoorAtlas that provides a geomagnetic indoor positioning technology to locate yourself inside buildings. Therefore we are going to use the IndoorAtlas technology for the indoor positioning and routing, which has proved to be robust and very accurate. It provides a free and open source technology to develop any kind of application related with indoor location.

2.3 Componens of a Smartphone

Smartphones and other portable electronic devices have become very popular in recent years. At the present time, most cellphones of the market already have a really large number of sensors that increase the possibilities of the device (Van Dalen and Oncala, 2013).

Here we present a brief list of sensors that modern cellphones have (see figure 5):

– Accelerometer

  The accelerometer measures proper acceleration which is the physical acceleration experienced by an object. This sensor is used in many applications like getting the speed of the device or counting the steps of the user.

– Gyroscope

  A gyroscope is useful for measuring the orientation of the device. Unlike the original gyroscope made of a spinning disk in which axis of rotation is free to assume any orientation by itself, smartphones have a electronic microchip that gets this information.

– Magnetometer

  This sensor is used for measuring the Earth's magnetic field and in geophysical surveys to detect magnetic anomalies.

– Proximity sensor

  As the name implies, a proximity sensor is able to detect the presence of nearby objects without any physical contact. In mobile devices this sensor is typically used to detect accidental touchscreen taps when held to the ear during a call. Therefore this sensor has only digital measure on smartphones.

– Light sensor

  A phone's light sensor measures how bright the ambient light is. It is commonly used to adjust the display's brightness automatically based on the light intensity of the environment.

- Barometer

  The recent smartphone models have a embedded barometer that can measure atmospheric pressure to determine how high the device is above the sea level to improve the GPS accuracy.

- Touchscreen

  The touchscreen is not just a sensor, it is an input device that allows the user to give input or control the smartphone through touching this screen.

- GPS

  Introduced in Section 2.1, this system is already installed in almost all the current cellphones providing a global location of the device.

- WiFi

  WiFi is a technology that allows electronic devices to connect to a WLAN network. Most modern smartphones have this technology allowing an internet connection through access points.

- Bluetooth

  Bluetooth is a wireless technology standard that connects devices together over a certain distance for exchanging data.

- GSM/CDMA Cell

  Every cellphone uses telephony base stations and antennas to make possible the communication between mobile phones. The most important concepts of this technology are the GSM standard and the CDMA protocols.

  The Global System for Mobile Communications or GSM standard was developed by the European Telecommunications Standards Institute (ETSI) to describe the protocols for second-generation (2G) digital cellular networks used by mobile phones. This digital mobile telephony system is widely used in Europe an many parts of the world (Mouly et al., 1992).

  The Code-Division Multiple Access (CDMA) refers to protocols used in 4G wireless communications. It is a form of multiplexing which allows several signals to occupy a single transmission channel, optimizing the use of available bandwidth (Lee and Miller, 1998).

- NFC

  Near Field Communication or NFC is a set of communication protocols that link two electronic devices (one of them is usually a smartphone or tablet) to establish a contactless communication. This system uses electromagnetic radio waves to send information without needing to touch the devices together or go through multiple steps setting up a connection (Curran et al., 2012).

As more mobile phone manufacturers start to include NFC chips in their devices, the applications of this technology will increase. However there are already many NFC applications like new marketing methods or new payment techniques.

– Front and back cameras

A camera is an optical instrument for recording and capturing images. Nowadays every mobile phone has at least one camera that can be used not only for image purposes but also for augmented reality and social applications.

Figure 5: Sensors of a current smartphone. (Source: http://smartphoneworld.me)

The most relevant application for indoor-outdoor detection is the preliminary research work done in IODetector (Zhou et al., 2012). This study presents a lightweight sensing service suitable for smartphones that mainly uses the light sensor, the magnetometer and cell tower signals. We are going to start from that to develop a more robust detection.

2.4 Light Sensor

As mentioned in the previous section, the light sensor measures the light intensity (Lux) on mobile phones. Current research on indoor-outdoor detection are based on this value because the sunlight outdoors gives a really high measurement compared to the artificial light that we used to use in indoor environments.

Therefore the light intensity inside buildings is much lower than in outdoor environments. The biggest reason for this phenomenon is that the light intensity incoming from the Sun within the visible spectrum is commonly higher than the incident light intensity from ordinary light sources (Zhou et al., 2012).

Figure 6 shows light intensity values in indoor, semi-outdoors and outdoor environments during different time points through the day.

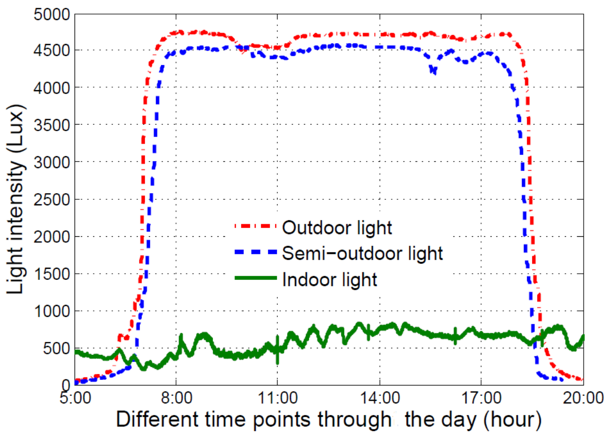

Figure 6: Outdoor and indoor light variation throughout the whole day (Zhou et al., 2012).

The problem of this measurement is that only works during daylight. Thus, to build a more effective detection method we can only take into account the light intensity during the day.

2.4.1 Sunrise-Sunset Algorithm

In order to apply the light intensity values to the detection, first we need to know if it is day or night. This can be easily done by knowing at which time is the sunrise and sunset. The sunrise-sunset equation can be used to estimate the time of sunrise and sunset for any solar declination and latitude (Meeus, 1991).

$$\cos w_0 = -\tan\phi \times \tan\delta,$$

where:

w_0 is the hour angle at either sunrise (when negative value is taken) or sunset (when positive value is taken),

ϕ is the latitude of the observer on the Earth, and

δ is the Sun's declination.

On a mobile device we cannot get directly the declination and we need the local sunrise and sunset time instead of the hour angle to compare it to current time. For this reason is better to use the sunset-sunrise algorithm (Williams). This algorithm needs the following data as inputs: time, day, year, local offset time, latitude, longitude and zenith. All these parameters can be taken from the smartphone except the zenith.

The Sun's zenith for sunrise and sunset can be taken by choosing one of the following values:

- Official: 90 degrees.
- Civil: 96 degrees.
- Nautical: 102 degrees.
- Astronomical: 108 degrees.

After obtaining the input parameters, the sunrise-sunset algorithm follows these steps to get the local time for the sunrise and sunset:

1 Calculate the day of the year.

$$N1 = floor(275 \times month/9),$$
$$N2 = floor((month + 9)/12),$$
$$N3 = (1 + floor((year - 4 \times floor(year/4) + 2)/3)),$$
$$N = N1 - (N2 \times N3) + day - 30.$$

2 Convert the longitude to hour value and calculate an approximate time,

$$lngHour = longitude/15.$$

For the rising time:

$$t = N + ((6 - lngHour)/24).$$

For the setting time:

$$t = N + ((18 - lngHour)/24).$$

3 Calculate the Sun's mean anomaly:

$$M = (0.9856 \times t) - 3.289.$$

4 Calculate the Sun's true longitude:

$$L = M + (1.916 \times \sin M) + (0.020 \times \sin(2 \times M)) + 282.634.$$

5 Calculate the Sun's right ascension:

$$RA = \arctan(0.91764 \times \tan L).$$

6 Right ascension value needs to be in the same quadrant as L:

$$Lquadrant = (floor(L/90)) \times 90,$$
$$RAquadrant = (floor(RA/90)) \times 90,$$
$$RA = RA + (Lquadrant - RAquadrant).$$

7 Right ascension value needs to be converted into hours:

$$RA = RA/15.$$

8 Calculate the Sun's declination:

$$sinDec = 0.39782 \times \sin L,$$
$$cosDec = \cos(\arcsin(sinDec)).$$

9 Calculate the Sun's local hour angle

$$cosH = (\cos(zenith) - (sinDec \times \sin(latitude)))/(cosDec \times \cos(latitude)),$$

if ($cosH > 1$) the sun never rises on this location (on the specified date),
if ($cosH < -1$) the sun never sets on this location (on the specified date).

10 Finish calculating H and convert into hours.
For the rising time:

$$H = 360 - \arccos(cosH).$$

For the setting time:

$$H = \arccos(cosH),$$
$$H = H/15.$$

11 Calculate local mean time of rising/setting.

$$T = H + RA - (0.06571 \times t) - 6.622.$$

12 Adjust back to UTC.

$$UT = T - lngHour.$$

13 Convert UT value to local time zone of latitude/longitude.

$$localT = UT + localOffset.$$

Thanks to the algorithm above we can calculate the rising and setting local time of the Sun providing a manner to know if is day or night.

2.5 Magnetometer

The magnetometer of a cellphone is used to measure the magnetic field of the Earth (Chung et al., 2011). This can be used for indoor-outdoor detection because the steel structures and electric sources of indoor environments disturb this value (Rowe et al., 2009). The disturbance of the Earth's magnetic field inside buildings can be utilized as fingerprints for the indoor localization as explained in Section 2.3. The IODetector research mentioned before (Zhou et al., 2012) applies the magnetic component to the indoor-outdoor detection by inspecting the variance of magnetic field intensity measured in a specific time period. The figure below plots a experiment to determine the magnetic field strength and its variance.

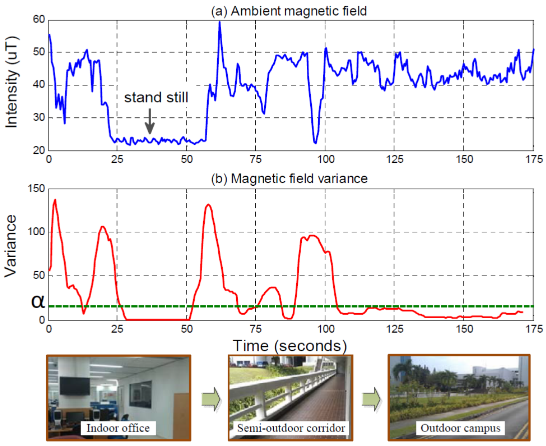

Figure 7: Variation of magnetic field strength (Zhou et al., 2012).

In outdoor environments the variation is really low because the magnetic field strength does not change as there are no magnetic disturbances. On the other hand inside buildings the variation of the magnetic field strength is higher than outdoors due to the metal structure and electrical components of the buildings that disturb the Earth's magnetic field intensity. Furthermore, the experiment shows that in indoor environments, the variation turns to zero if you stand still so it is critical to also detect if you are moving to correct this fact.

To solve this, the IODetector research (Zhou et al., 2012) uses the accelerometer to discover if you are moving or stand still. With this fix, this system uses a threshold to detect if you are indoors or outdoors. If the magnetic field variance in a time window of 10 seconds is lower than 18 it is probable that you are outdoors, otherwise you are indoors.

This is good approach to use the magnetometer of mobile phones to detect if you are indoors or outdoors. However this component of the IODetector has the lowest accuracy (Radu et al., 2014).

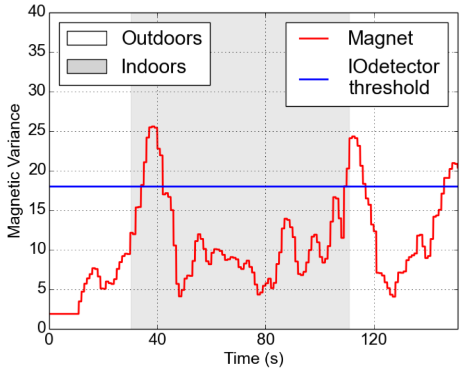

Figure 8: Variation of magnetic field strength with given threshold (Zhou et al., 2012).

Figure (Figure 8) disclose that this tool is not as robust as we expected so we need better ways to detect the indoor-outdoor transition. As mentioned, it is a fact that the magnetic field intensity is more constant outdoors than indoors. In outdoor environments the magnetometer measures the Earth's magnetic field in the user's location while in indoor environments this value is distorted.

Thus we came up with the idea of comparing the magnetic field strength measured by the magnetometer with the value of the Earth's magnetic field in that current location. If the difference between these values is low enough, the user is probably outdoors because the magnetometer measures a value similar to the Earth's magnetic field value. Otherwise the user is indoors because the magnetometer compute a value that does not match with the Earth's magnetic field intensity in that location due to disturbances caused by metal objects and electrical devices.

There are several ways to calculate the Earth's magnetic field intensity on a current location. The more interesting one is the dipole model of the Earth's magnetic field which is going to be explained in the following section.

2.5.1 Dipole model of the Earth's magnetic field

The dipole model of the Earth's magnetic field is a first order approximation of the Earth's magnetic field. This model is particularly inaccurate at high L-shells (particular set of planetary magnetic field lines) because of the effects of the interplanetary magnetic field and the solar wind but is a good and simple approximation for L-shells lower than three.

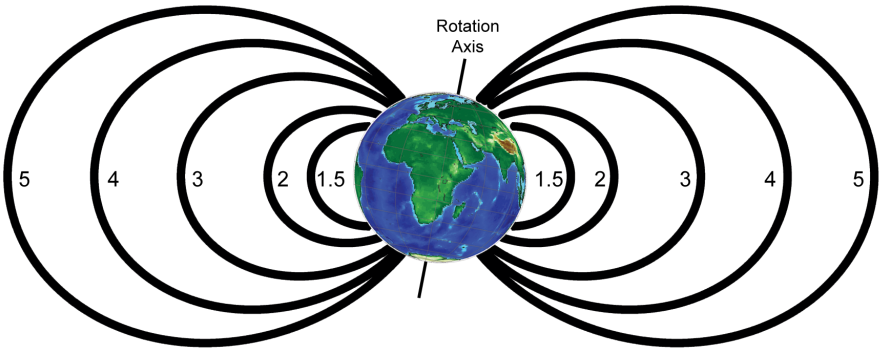

Figure 9: L-values using a dipole model of the Earth's magnetic field. (Source: www.wikipedia.com)

Now we present the equations that describe the dipole magnetic field model (Walt, 2005). The magnitude for the magnetic field (B) is the combination of the radial ($\mathbf{B_r}$) and azimuthal ($\mathbf{B}_\theta$) fields that can be described as follows:

$$\mathbf{B_r} = -2B_0 \times \left(\frac{R_E}{r}\right)^3 \times \sin\lambda,$$

$$\mathbf{B}_\theta = B_0 \times \left(\frac{R_E}{r}\right)^3 \times \cos\lambda,$$

$$|\mathbf{B}| = B_0 \times \left(\frac{R_E}{r}\right)^3 \times \sqrt{1 + 3\sin^2\lambda},$$

where:

- $\mathbf{B_0}$ is the mean value of the magnetic field at the magnetic equator on the Earth's surface. Commonly $B_0 = 3.12 \times 10^{-5} \times T$,
- $\mathbf{R_E}$ is the mean radius of the Earth (6370 km approximately),
- λ is the latitude measured northwards from the equator,
- $\mathbf{r}$ is the radial distance from the center of the Earth (same units as R_E).

To get the magnetic field in the current location of the user, we need to know the latitude (λ) which is an available input in smartphones, as well as the radial distance (r) which can be calculated based on the latitude as follows (Moritz, 1980):

$$\mathbf{r} = \sqrt{\frac{(a^2\cos\lambda)^2 + (b^2\sin\lambda)^2}{(a\cos\lambda)^2 + (b\sin\lambda)^2}},$$

Where a is the equatorial radius (6,378.1370 km) and b is the polar radius (6,356.7523 km).

2.6 Cell Strength

In addition to the magnetic field and the light intensity, the IODetector (Zhou et al., 2012) focuses on detecting the change of cellular signal strengths using the GSM standard (see Section 2.3). The Received cellular Signal Strength (RSS) is used to measure the signal intensity. The absolute value of the RSS shows that inside buildings the signal strength is significantly lower than in outdoor environments due to the walls and other structures that may block the cellular signal as shown in Figure 10.

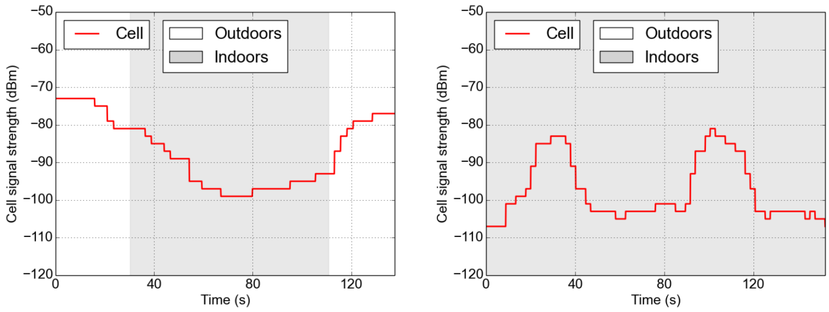

Figure 10: Left: Cell signal strength at transition. Right: Cell signal variations indoor (Zhou et al., 2012).

This value is very inaccurate because it is directly related to the movement of the user. For instance, if you are inside and you move in the proximity of a windows the RSS will increase (Figure 10). Also in outdoor environments if you hide behind a tree for example, the RSS will decrease.

To exploit the RSS more accurately the IODetector uses the RSS variation within a short period of time to indicate the context transition. This is useful to detect the indoor-outdoor transitions but it cannot detect the state when the user is static (Radu et al., 2014). Therefore a robust detection cannot be based entirely on the RSS.

2.7 GPS

The GPS, available in practically any current cellphone, allows the user to know its location with a certain accuracy. GPS signals depends directly on how well the smartphone "sees" the satellites placed in the orbit around the Earth. Outdoors the mobile phone can "see" the GPS satellites easier than indoors where the sky is obscured by ceiling and walls (Ravindranath et al., 2011).

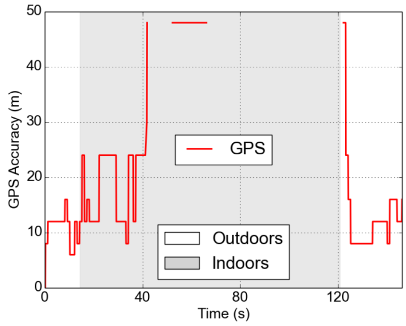

Figure 11: GPS accuracy at indoor-outdoor transitions (Zhou et al., 2012).

Recent studies use the estimated accuracy of GPS localization to detect the indoor-outdoor transitions as shown in Figure 11 (Radu et al., 2014). However it is not a good method to detect the indoor-outdoor state by its own because the GPS can sometimes get a satellite fix indoors and give a wrong accuracy value, for instance when the user come close to a window or door. There are still other ways to use the GPS system to know whether you are indoors or outdoors like counting the number of satellites available and measuring the signal strength of the satellites.

2.7.1 Number of Satellites

Like the accuracy of the GPS, the number of visible satellites is a value that provides useful information to detect indoor and outdoor environments.

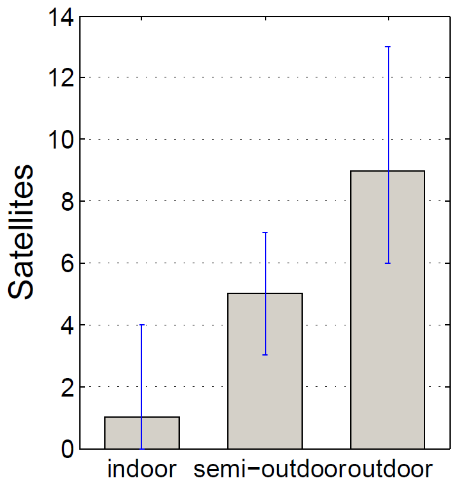

Figure 12: Number of visible satellites in indoor, semi-outdoor and outdoor environments (Zhou et al., 2012).

In Figure 12 we find that the number of observed satellites varies from indoors to outdoors. Inside buildings, mobile phones normally receive two or less GPS signals, though the mobile phones can sometimes receive more GPS signals near windows and doors. In outdoor environments cellphones normally receive signals from more than 6 GPS satellites even in cloudy and rainy days (Zhou et al., 2012).

2.7.2 NMEA Data

In the world of GPS, NMEA (National Marine Electronics Associationis [1]) is a standard data format supported by all GPS manufacturers. The NMEA data provides a combination of electrical and data specifications for communication between marine equipments such as sonars, gyrocompass, GPS receivers and many other instruments. The standard, as implemented in GPS receivers, has been made sufficiently general that can be found in several non-marine applications around the world.

To transmit the data, the standard uses the concept of "talkers" and "listeners" to send and receive simple plain text with the characters coded using seven-bit ASCII (American Standard Code for Information Interchange) (Langley, 1995). The data format consists of a sentence with a maximum of 82 characters, always beginning with the starting delimiter "$" and ending with a carriage return(<CR>) or line feed (<LF>) delimiters. The data is contained in a single line with data items separated by commas. The first word of the sentence is called the data type which defines the interpretation of the rest of the sentence. Each data type has its unique interpretation and is defined in the NMEA standard (see http://www.gpsinformation.org/dale/nmea.htm).

There are several sentences in the NMEA standard for all kind of devices that may be used in a Marine environment. The messages that may be useful for GPS receivers are the ones that start with GP. Within these messages, the more relevant sentences are presented below with a brief description (see http://aprs.gids.nl/nmea).

- $GPBOD - Bearing, Origin to Destination.
- $GPBWC - Bearing using Great Circle Waypoint.
- $GPGGA - Global positioning system fix information.
- $GPGLL - GPS Latitude and Longitude data.
- $GPGSA - GPS Active Satellites.
- $GPGSV - GPS, Satellites in View.
- $GPHDT - Heading, True.

[1] http://www.nmea.org

- $GPR00 - List of waypoints in currently active route.
- $GPRMA - Recommended minimum specific Loran-C data.
- $GPRMB - Recommended minimum navigation info for gps.
- $GPRMC - Recommended minimum specific GPS/Transit data.
- $GPRTE - Routes message.
- $GPTRF - Transit Fix Data.
- $GPSTN - Multiple Data ID.
- $GPVBW - Dual Ground / Water Speed.
- $GPVTG - Vector Track and speed over the Ground.
- $GPWPL - Waypoint location information.
- $GPXTE - Measured cross-track error.
- $GPZDA - Date and Time.

For the indoor-outdoor detection, the most important sentence is the satellites in view (GPGSV) because it has all the information necessary. Each sentence can only provide data for up to 4 satellites and thus there might be need 3 sentences for the full information. The example below will clarify the information within this type of NMEA message.

$GPGSV,3,1,11,03,03,111,00,04,15,270,00,06,01,010,00,13,06,292,00*74

$GPGSV,3,2,11,14,25,170,00,16,57,208,39,18,67,296,40,19,40,246,00*74

$GPGSV,3,3,11,22,42,067,42,24,14,311,43,27,05,244,00„„*4D

The following list explains the characters of each sentence according to their position, with the corresponding example values in parenthesis.

1 : Total number of messages of this type in this cycle (3, 3, 3).

2 : Message number (1, 2, 3).

3 : Total number of SVs in view (11, 11, 11).

4 : Satellite PRN number (03, 12, 22).

5 : Elevation in degrees, 90 maximum (03, 25, 42).

6 : Azimuth, degrees from true north, 000 to 359 (111, 170, 067).

7 : SNR (Signal-Noise Ratio), 00-99 dB and null when not tracking (00, 00, 42).

8-11 : Information about second satellite, same as field 4-7.

12-15 : Information about third satellite, same as field 4-7.

16-19 : Information about fourth satellite, same as field 4-7

The field called SNR (Signal to Noise Ratio) is a value that gives information about the strength of the signal. This ratio compares the level of the satellite GPS strength to the level of background noise. The higher value of SNR the better accuracy the satellite has (Hetet, 2000).

Returning to the aim of the thesis, SNR value is really helpful for the indoor-outdoor detection because in outdoor environments the SNR value is higher than inside buildings where the GPS does not work so well (see Figure 13).

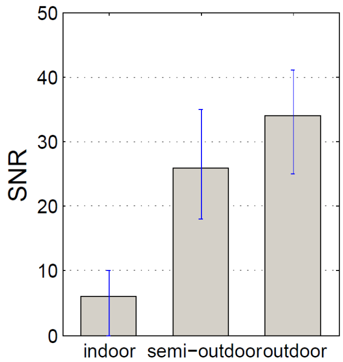

Figure 13: SNR in indoor, semi-outdoor and outdoor environments (Zhou et al., 2012).

We will look into this with more detail in Section 3.3.3.

2.8 Bayesian Networks

A Bayesian network or *belief network* is a probabilistic graphical model that is used to represent a set of random variables and their conditional dependencies. The random variables are the nodes of the diagram while the edges between nodes are the conditional dependencies as shown in Figure 14 (Charniak, 1991).

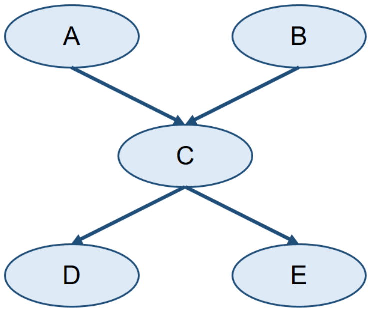

Figure 14: Simple Bayesian network.

Bayesian networks use a structure called direct acyclic graph (in terms of graphical models) which allows an easy understanding of the maths behind it. The arrows of the edges point from parent nodes to child nodes giving hierarchical information to the diagram. This means that the child nodes depends on the value taken by its parent nodes (Ben-Gal, 2007).

Aside from the graph, we need to quantify the conditional probability distribution at each node of the Bayesian network which meet the Markov property (it only depends on its parent). These probabilistic distributions must take into account all possible combinations of their parents and can be continuous or discontinuous. In the following example we assume that the states of the nodes can only be true (A) or false ($\neg A$).

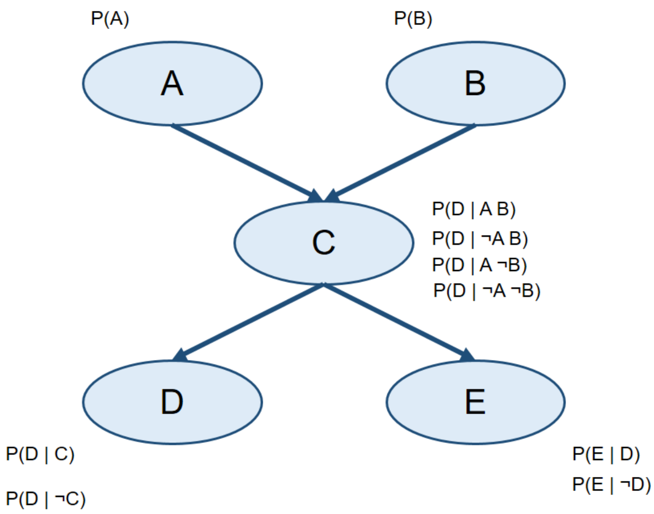

Figure 15: Bayesian network, known probabilities.

Bayesian networks are based on the Bayes' theorem which describes the probability of an event, based on conditions that might be related to the event

$$P(X|Y) = \frac{P(X)P(Y|X)}{P(Y)},$$

where:

$P(X)$ and $P(Y)$ are the probabilities of events X and Y without regarding each other (with $P(B) \neq 0$),

$P(X|Y)$ is the conditional probability observing event X given Y,

$P(Y|X)$ is the conditional probability observing event Y given X,

In the Bayesian network context, it is easier to write the Bayes' theorem as:

$$P(X|Y) = \frac{P(X)P(Y|X)}{P(Y|X)P(X) + P(Y|\neg X)P(\neg X)}$$

Where:

$P(X)$ is the initial probability or prior probability of the node X.

$P(\neg X)$ is $P(\neg X) = 1 - P(X)$.

$P(Y|X)$ is the conditional probability or likelihood of the node Y given X is true.

$P(Y|\neg X)$ is the conditional probability or likelihood of the node Y given X is false.

$P(X|Y)$ is the final probability for the node X after taking into account the influence of Y.

The example on Figure 15 also help us to show the three different connections that may appear on a Bayesian networks. The first type (Figure 16 left) is the linear path where a node has a single parent node, the second (Figure 16 middle) is the diverging path where a node has more than one child node and the third one (Figure 16 right) is the converging path where a child node has one or more parent nodes (Charniak, 1991).

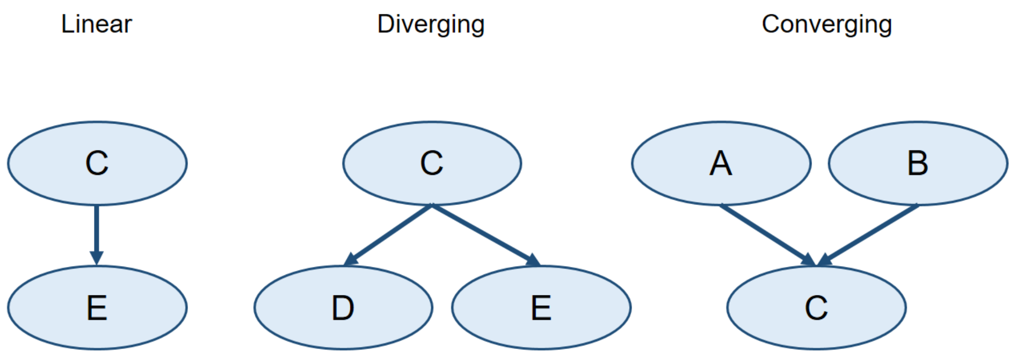

Figure 16: Kinds of connections in Bayesian networks.

Each Bayesian network has a unique joint probability function which contains all the information of how the Bayesian network is modelled. For the example presented above, the joint probability function is the following (Pearl and Russell, 1998):

$$P(A, B, C, D, E) = P(A)P(B)P(C|A, B)P(D|C)P(E|C)$$

Now we can calculate the probability of any node in terms of the conditional probabilities. For example, the probability that A is true, given D and E as true.

$$P(A|D, E) = \frac{P(A = T, D = T, E = T)}{P(D = T, E = T)}$$

$$P(A|D, E) = \frac{\sum_{B,C} P(A = T, B, C, D = T, E = T)}{\sum_{A,B,C} P(A, B, C, D = T, E = T)}$$

$$P(A|D, E) = \frac{\sum_{B,C} P(A = T)P(B)P(C|A = T, B)P(D = T|C)P(E = T|C)}{\sum_{A,B,C} P(A)P(B)P(C|A, B)P(D = T|C)P(E = T|C)}$$

For our project, Bayesian networks will allow us to model the robust indoor-outdoor detection system explained in Section 3.4.

3 Research Material and Methods

In this thesis we have to distinguish between the research study we have made related to the indoor-outdoor detection and its implementation in an Android application for the indoor-outdoor guidance. In this section we are going to explain the research part where we are going to explain how we have made a robust algorithm to know whether you are indoors or outdoors.

3.1 Resources

The aim of the research is to develop a robust algorithm that detects if you are indoors or outdoors using a current smartphone. The first step to build it up is to gather information of the sensors available on a common mobile phone.

Figure 17: Sony Ericsson M2. (Source: www.sonymobile.com)

The cellphone used for that purpose is a *Sony Ericsson M2* (Figure 17) and here we present a brief list of its specifications.

- Network technology: GSM/HSPA/LTE
- Platform: Android Operative System 4.4.4 (Kit-Kat).
- Memory:

  - Internal: 5 Gb and 1 Gb RAM (Random Access Memory).
  - External: 8 Gb microSD card.

- Communications:
  - WLAN
  - GPS
  - NFC
  - Bluetooth

- Sensors:
  - Ambient light sensor
  - Magnetometer
  - Proximity sensor
  - Accelerometer

Once we have selected the device, it is time to get as much information of the sensors as we can in indoor-outdoor transitions.

To gather that information we have used a free application called AndroSensor, (available in the Play Store), that can record the measurements of all the sensors over time.

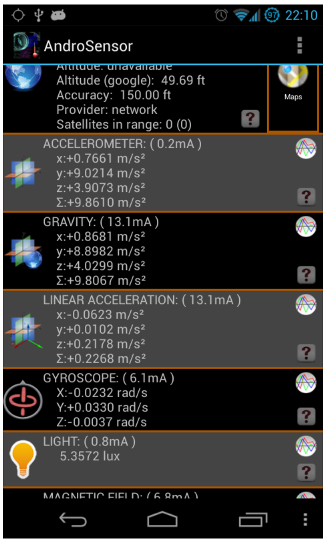

Figure 18: AndroSensor screenshot.

This dataset is stored in the mobile phone and afterwards transferred to the laptop for data processing.

The data processing was done with Matlab, a well known computing environment that provides a useful tool to analyse and solve mathematical problems. This software also allows the user to import data and plot any kind of functions.

We have collected 20 datasets in the months of January and February. Fourteen of them with the GPS enabled and six without. Furthermore we are going to explain how we have developed the algorithm on the basis of the following experiments.

3.2 Experiments without GPS

The six samples without the GPS were made to get the preliminary values that may help us to detect the indoor-outdoor transition. For instance, Figure 19 shows the dataset gathered on the 8-January-2016 at 13:00 approximately.

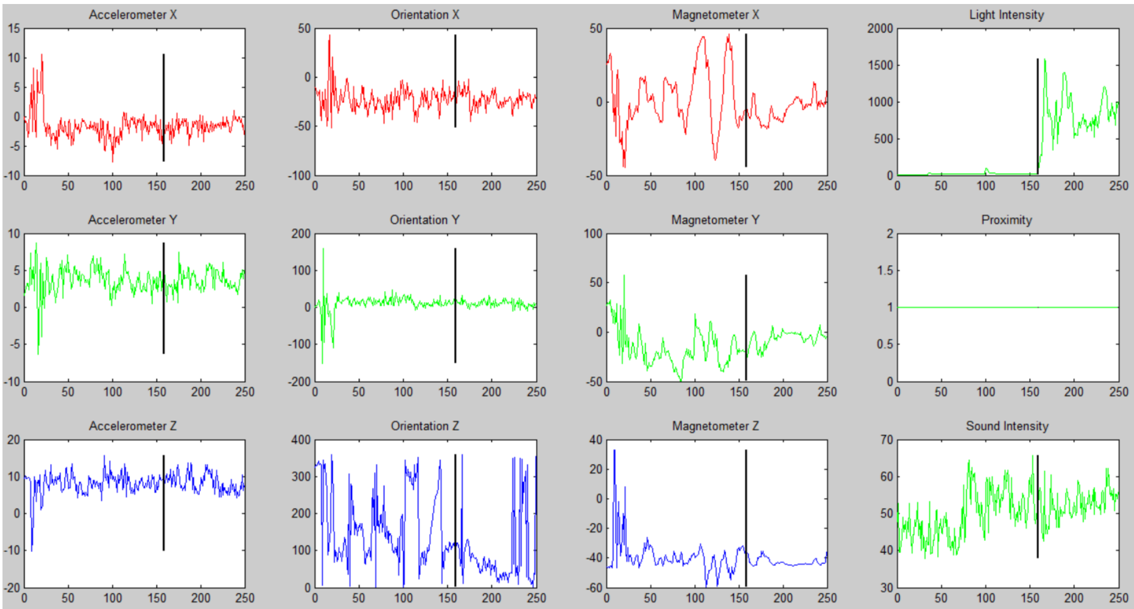

Figure 19: 8-January-2016 sample.

The first column illustrates the accelerometer values for the X, Y and Z axis. The second column refers to the orientation in X, Y and Z axis. The third columns shows to the magnetic field values for the X, Y and Z axis. And the last column shows the measurements of the light sensor (first row), the proximity sensor (second row) and the microphone (third row). The black vertical line represents in this case the change over indoors and outdoors. As the rest of the samples without using the GPS, the relevant values for detecting the state (indoor or outdoor) are the light intensity and the magnetic field.

3.2.1 Light Intensity

Focusing on the light intensity of the previous example, in Figure 20 we can observe that the light intensity in outdoor environments is higher than indoors even in cloudy days as we said in Section 2.4.

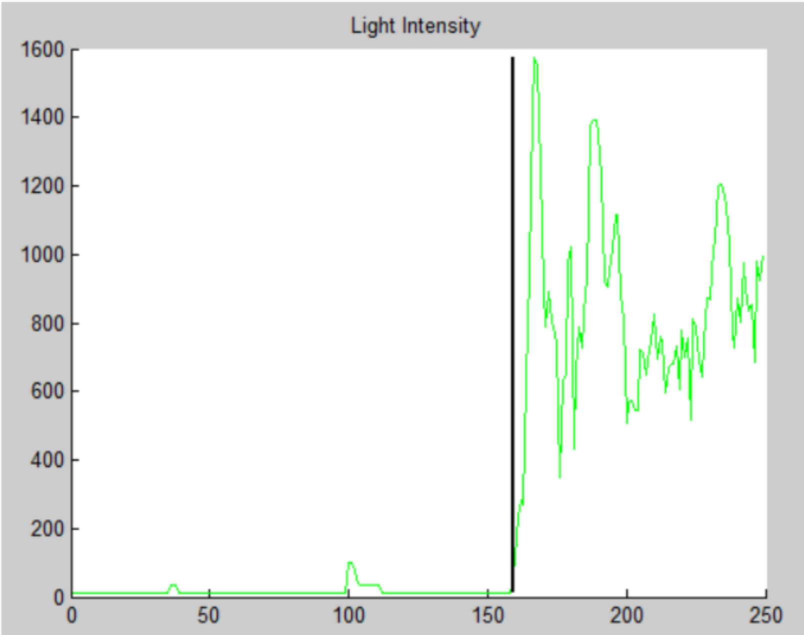

Figure 20: 8-January-2016 Light Intensity value.

The result of the experiments are collected in the following table where the minimum and maximum values of the light intensity are taken in three different weathers: sunny, partly cloudy and cloudy.

Weather	Light Intensity Range Outdoors	Light Intensity Range Indoors
Sunny	2500 - 100000 Lux	150 - 1750 Lux
Partly cloudy	1250 - 7500 Lux	15 - 500 Lux
Cloudy	750 - 3600 Lux	5 - 250 Lux

Therefore if the light sensor measures more than 1750 Lux the user is probably outdoors while if the light intensity is less than 750 Lux the user is probably indoors.

However these values may vary considering some external factors like light sources inside building or shadow spots in outdoor environments. For that reason we are going to place the absolutely certainty of being outdoors in 2500 Lux and the absolutely certainty of being indoors in 0 Lux.

3.2.2 Magnetic Field

To simplify the magnetometer measurements, we can compute the absolute value of the magnetic field (M) instead of each component (M_X, M_Y and M_Z) measured by the magnetometer.

$$M = \sqrt{M_X^2 + M_Y^2 + M_Z^2}$$

Thus, the absolute magnetic field value of the previous example is:

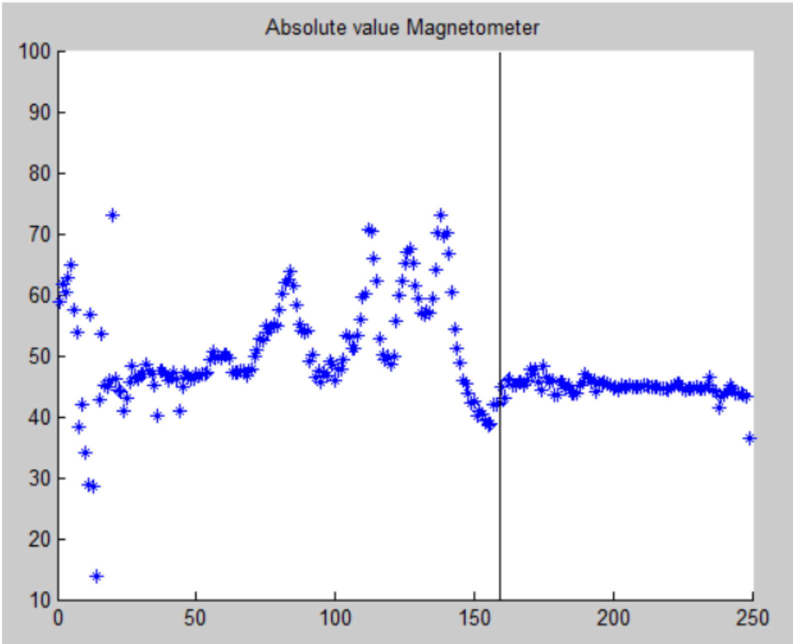

Figure 21: 8-January-2016 Absolute value of the magnetic field.

In Figure 21 above we can observe that the magnetic field indoors changes more than in outdoor environments caused by the metallic structures of the buildings that interfere with the measurements of the magnetometer. Nevertheless the value of the the variation of the magnetic field cannot be used to detect the state of the user because if the device is still the variation of the magnetic field will be close to 0 regardless if you are indoors or outdoors.

To use the magnetic field factor more robustly we decided to compare it against the Earth's magnetic field in that location. Without the disturb of metallic structures the magnetic field value should match with the Earth's magnetic field in that location. Otherwise we can measure a difference between the Earth's magnetic field in the current location and the value given by the magnetometer.

We have used the dipole model of the Earth's magnetic field explained in Section 2.5.1 to get that value using the location given by the GPS.

3.3 Experiments with the GPS enabled

We have made fourteen experiments with the GPS enabled to explore other values that may change between indoors and outdoors beyond the light intensity and the magnetic field.

We are going to use the sample made on the 14-January-2016 (Figure 22) at 13:10 as an example to illustrate all the values that turn from inside buildings to outdoor environments.

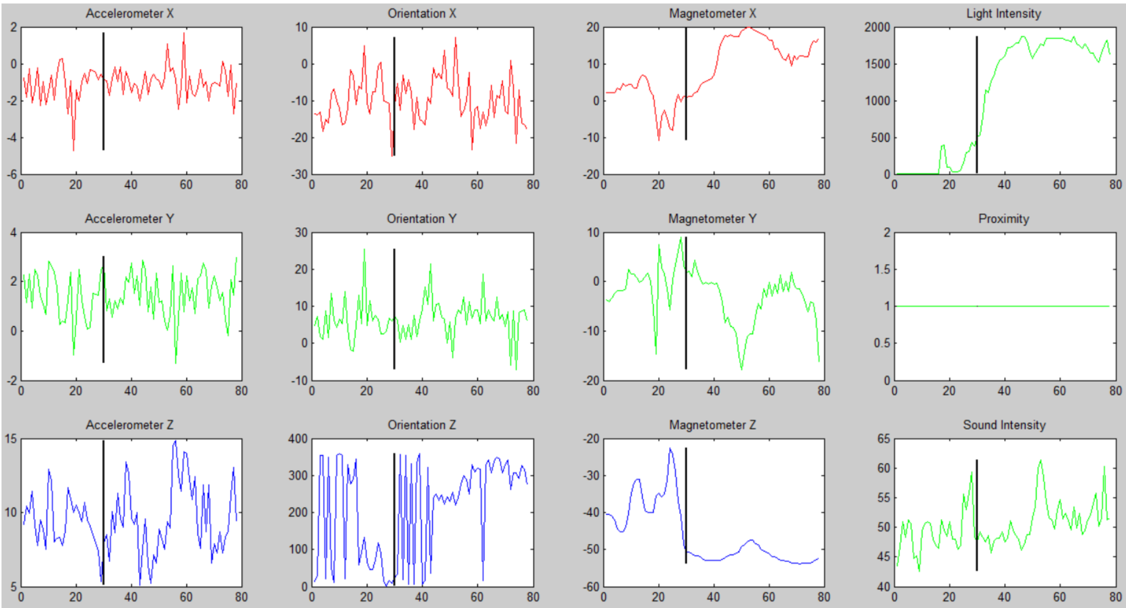

Figure 22: 14-January-2016 sample with the GPS enabled.

Again the light intensity and the magnetometer are the critical values that change in the indoor-outdoor transition but in the following paragraphs we are going to see other aspects that are influenced by the state of the device (indoor or outdoor).

First of all we have taken a look at the difference between the Earth's magnetic field in the current location of the user and the magnetic field measured by the magnetometer, as we have commented in the previous section.

Afterwards, we focused on the availability of the satellites (see Section 2.7) by counting the number of the satellites the smartphone can see and rating the signal to noise ratio.

3.3.1 Magnetic Field Difference

We used the dipole model of the Earth's magnetic field to get the hypothetical value of the magnetic field on the Globe's surface accordingly to the global position of the user (section 2.5.1). The dipole model of the Earth's magnetic field is a first order approximation that allows us to get a value of the magnetic field on the basis of the current latitude.

In Figure 23 we can observe in red color the value of the magnetic field calculated with the dipole model. This value is approximately constant during the experiments due to the magnetic field is determined with the current latitude of the user and that latitude does not change too much in the same region.

On the other hand, the blue color represents the absolute magnetic field measured by the magnetometer that changes more indoors than outdoors.

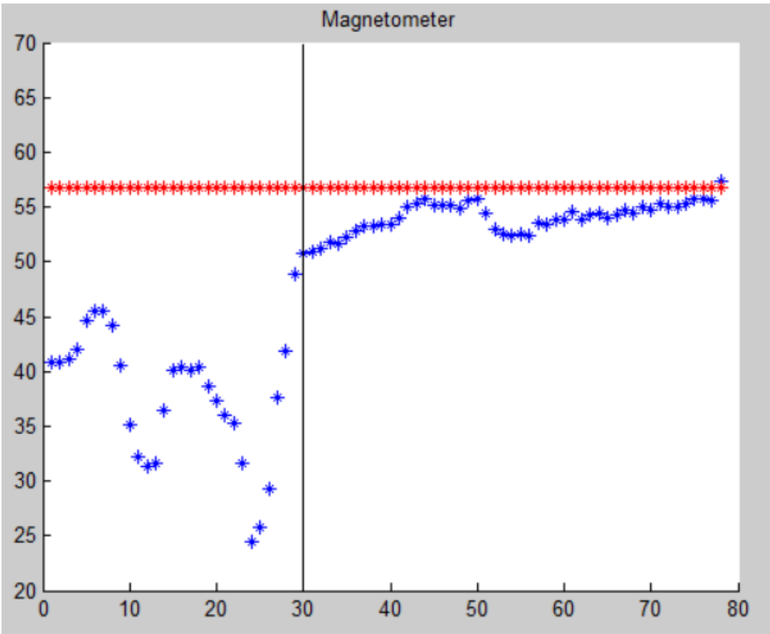

Figure 23: 14-January-2016 comparison between the absolute value of the magnetic field (blue) and the Earth's magnetic field at the user's location using the dipole model (red).

As we can see, the magnetic field measured by the magnetometer is really close to the hypothetical Earth's magnetic field outdoors while indoors that variation is higher because of the presence of metallic structures inside buildings.

The following table shows range of the magnetic field difference between the Earth's magnetic field and the magnetic field measured by the magnetometer indoors and outdoors.

State	Minimum Magnetic Field Difference	Maximum Magnetic Field Difference
Indoors	10 T	60 T
Outdoors	0.5 T	10 T

Hence if the magnetic field difference is higher than 10 T the user is probably indoors, else the user is probably outdoors. Nonetheless the magnetic field difference outdoors may increase because of other metallic equipments and electric supplies. Besides the interferences of the magnetic field indoors may result a magnetic difference lower than 10 T so we have to take carefully the threshold.

3.3.2 Number of Satellites

Introduced in section 2.7.1 the number of satellites visible from the device is directly related to the state of the user. The more satellites detected by the smartphone the more probability the user has to be outside.

Therefore we have gathered the number of satellites available during the indoor-outdoor transitions of the samples to prove this fact. As before, we have used the dataset of 14-January-2016 (Figure 24) to clarify it.

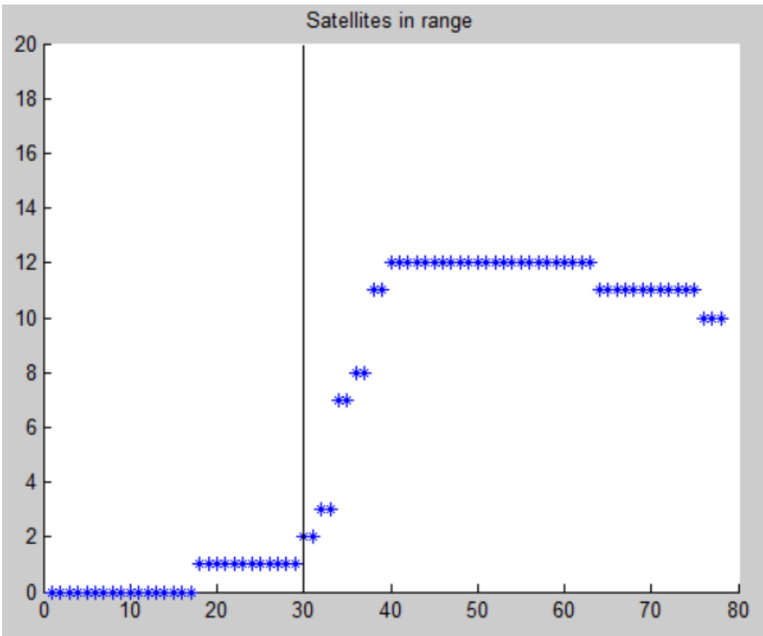

Figure 24: 14-January-2016 Number of satellites available.

We noticed that the GPS takes a while to detect all the satellites available when you go from indoors to outdoors (as the example). Also the GPS takes a little time to discard the unseen satellites walking into buildings from outdoor environments.

The number of satellites is a robust parameter to detect the state of the user taking into account the delay of the value measured. Taking into consideration all the experiments, we made a table to set the range of values indoors and outdoors.

State	Minimum Number of Satellites	Maximum Number of Satellites
Indoors	0	7
Outdoors	8	21

Although the number of satellites measured indoors can be higher than 7 next to windows or doors, the aim of the final algorithm is to detect the indoor-outdoor transitions. Thus, the results illustrate that the number of satellites measured indoors is lower than 7 while outdoors is higher than 8.

3.3.3 Signal to Noise Ratio

The Signal to Noise Ratio (SNR) in the GPS context, gives information about the quality of the signal emitted from the satellites to the device. In Section 2.7.2 we explained that this ratio, obtained from the NMEA data, is a value that measures how good is the signal compared to the external noise.

To collect this value we have parsed the NMEA data received from the device as it is explained in the following example.

The next sentences are a sample file of the NMEA data randomly generated by the source Random NMEA Sentences Generator[2].

$\$GPRMC, 100311.518, V, 3454.931, N, 07902.499, W, 22.3, 3.48, 010616, , E44$
$\$GPGGA, 100312.518, 3454.931, N, 07902.499, W, 0, 00, , , M, , M, , 59$
$\$GPGLL, 3454.931, N, 07902.499, W, 100313.518, V38$
$\$GPGSA, A, 2, 07, 10, 01, 03, 10, 13, , , , , , , 0.1, 0.6, 0.83B$
$\$GPGSV, 2, 1, 06, 07, 41, 022, 86, 10, 07, 222, 18, 01, 74, 247, 63, 03, 05, 101, 3678$
$\$GPGSV, 2, 2, 06, 10, 62, 070, 44, 13, 19, 318, 717B$

The receiver gets one of these sentences at a time so the first step is to get only the sentences necessary to measure the SNR of the satellites. That sentences are the ones which start with $\$GPGSV$.

$\$GPGSV, 2, 1, 06, 07, 41, 022, 86, 10, 07, 222, 18, 01, 74, 247, 63, 03, 05, 101, 3678$
$\$GPGSV, 2, 2, 06, 10, 62, 070, 44, 13, 19, 318, 717B$

Each sentence has at most four satellites fixed so the sentence will give at most four values of the SNR that are placed in the positions 7, 11, 15 and 19 of the sentence as it is explained in Section 2.7.2.

[2]http://freenmea.net/emitter

$\$GPGSV, 2, 1, 06, 07, 41, 022, 86, 10, 07, 222, 18, 01, 74, 247, 63, 03, 05, 101, 3678$

Position 7: $SNR_1 = 86dB$
Position 11: $SNR_2 = 18dB$
Position 15: $SNR_3 = 63dB$
Position 19: $SNR_4 = 36dB$

$\$GPGSV, 2, 2, 06, 10, 62, 070, 44, 13, 19, 318, 717B$

Position 7: $SNR_1 = 44dB$
Position 11: $SNR_2 = 71dB$
Position 15: null
Position 19: null

Through all the $\$GPGSV$ sentences we are going to take the average of the SNR received like:

$$SNR = \frac{(SNR_1 + SNR_2 + SNR_3 + SNR_4)_1 + (SNR_1 + SNR_2)_2}{6}$$

$$SNR = \frac{(86 + 18 + 63 + 36) + (44 + 71)}{6} = 53dB$$

Once we know how the SNR is measured, we can proceed with the experiments and get the range of values indoors and outdoors. Figure 25 continues with the example in previous paragraphs measuring the SNR the 14-January-2016 from inside a building to an outdoor environment.

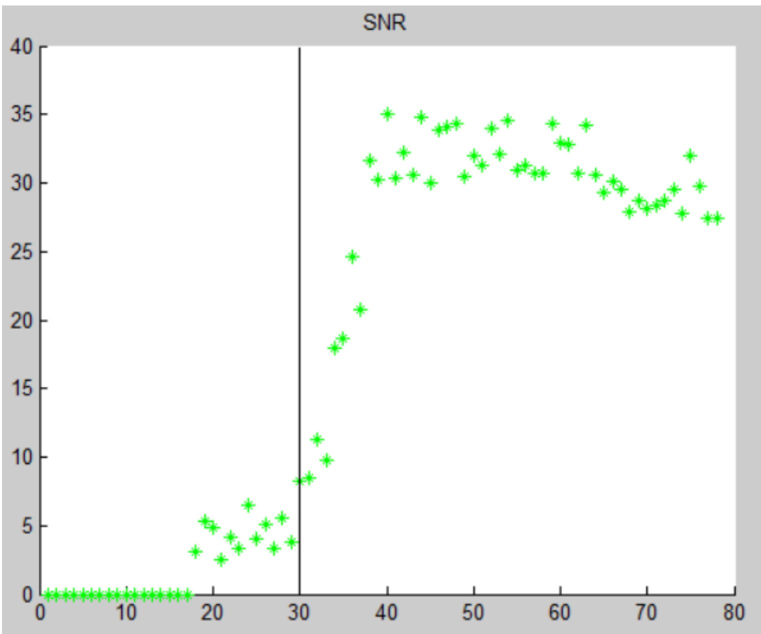

Figure 25: 14-January-2016 Signal to noise ratio.

As expected, the higher SNR the better the quality of the satellites signal and therefore the probability of being outdoors is higher. This phenomenon occur because inside buildings the receiver cannot see the satellites directly due to the walls and ceiling. On the other hand when the user is outdoors, the device can see better the satellites available because there are no obstacles.

The SNR range is from $0dB$ up to $99dB$ but smartphones often do not exceed the $50dB$. The next table shows the range of SNR values indoors and outdoors gathered from all the experiments.

State	Minimum SNR	Maximum SNR
Indoors	0 dB	20 dB
Outdoors	21 dB	36 dB

As we can see, with a SNR lower than $20dB$, the user is probably indoors while if it is higher than $21dB$ it is more likely that the user is outdoors. The SNR is a parameter for indoor-outdoor detection even more robust than the number of satellites because it is easier to set the transition threshold.

3.4 Bayesian Network

This part will describe how we used the Bayesian networks to develop a robust algorithm that detects whether you are indoors or outdoors. The mathematical background of Bayesian networks is explained in Section 2.8 and here we are going specify the parameter for this particular case.

The aim of our system is to get the probability of being outdoors depending on the parameters mentioned before which are the light intensity (L), the magnetic field difference (M), the number of satellites (N_{Sat}) and the signal to noise ratio of the satellites (Snr). We have chosen the probability of being outdoors ($P(Out)$) rather than the probability of being indoors ($P(In)$) randomly because $P(In) = 1 - P(Out)$ and both contain the same information. The network is represented in Figure 26.

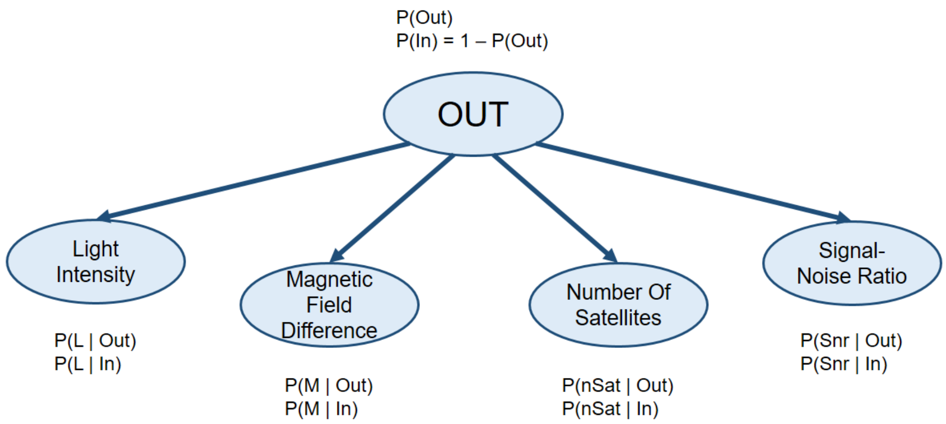

Figure 26: Bayesian network probabilities applied to the indoor-outdoor detection.

Following the steps to build our Bayesian network we need the probabilities of each parameter of being outdoors ($P(Y|Out)$) and of being not outdoors ($P(Y|\neg Out)$), that is the probability of being indoors ($P(Y|In)$). For that purpose we have to turn the measures of the parameters gathered in the experiments into probability distributions.

These probability distributions $p(y|Out)$ are defined in the interval $[a, b]$ and therefore the integral of the function from a to b should be one.

$$\int_a^b p(y|Out)dy = 1, \forall y \in [a, b]$$

The key now is to find the function $p(y|Out)$ and the interval $[a, b]$ for each parameter that satisfy the statement above accordingly with the results of the experiments.

3.4.1 Light Intensity Probability Distribution

Taking into account the experiments, the interval of values for the light intensity probability distribution is from 0 Lux to 2500 Lux. A value of 0 Lux means that the user is absolutely indoors while with 2500 Lux the user is certainly outdoors.

3.4.1.1 Light Intensity Probability of being Outdoors

The probability of being outdoors for the light intensity is high and almost constant with values close to 2500 Lux and decreases fast when the measurement of the light sensor is around 0 Lux. Thus the behaviour of the light intensity probability outdoors stands for a square root $p(y|Out) = \alpha\sqrt{y}$ where α is a parameter that assures the integral sums one,

$$\int_0^{2500} \alpha\sqrt{y}dy = 1, \forall y \in [0, 2500].$$

To obtain α we have simplified the measurement read by the light sensor dividing it by 100 so now the range is $[0, 25]$,

$$\int_0^{25} \alpha\sqrt{y}dy = 1, \forall y \in [0, 25].$$

Solving this integral we get an α value of $\alpha = 0.012$. The final probability distribution of being outdoors for the light intensity is represented in Figure 27,

$$p(y|Out) = 0.012\sqrt{y}, \forall y \in [0, 25].$$

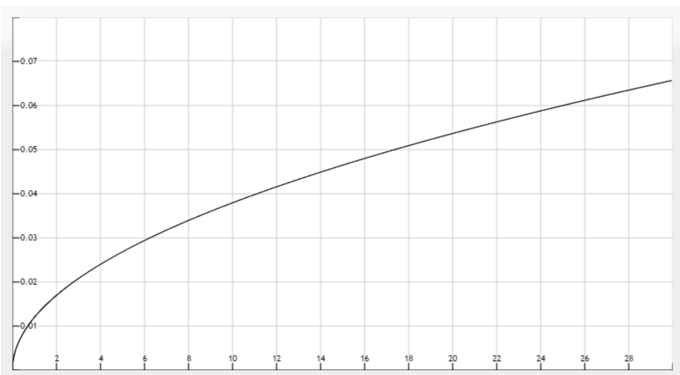

Figure 27: Light intensity probability distribution of being outdoors.

3.4.1.2 Light Intensity Probability of being Indoors

On the other hand the probability of being indoors for the light intensity should decrease fast from 0 to 2500 Lux. Again we have divided the measurements of the light sensor by 100 to simplify the function so the interval is $[0, 25]$. To match the behaviour we have used the function $p(y|In) = \alpha(y - 25)^2$ and again we have calculated α to set the integral to one.

$$\int_0^{25} \alpha(y - 25)^2 dy = 1, \forall y \in [0, 25]$$

The result is an $\alpha = 0.000192$ and therefore the probability distribution of being indoors is:

$$p(y|In) = 0.000192(y - 25)^2, \forall y \in [0, 25].$$

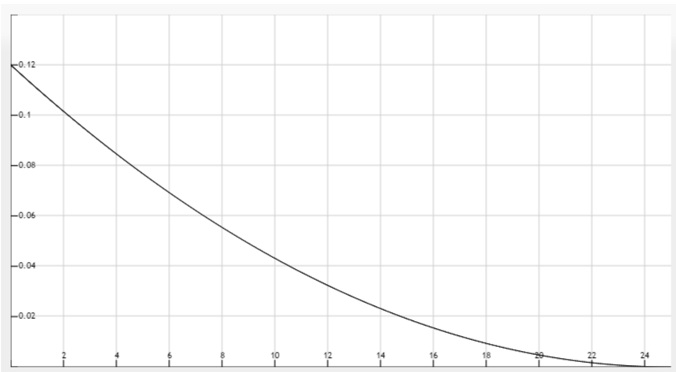

Figure 28: Light intensity probability distribution of being indoors.

3.4.2 Magnetic Field Difference Probability Distribution

The interval of the magnetic field difference values is set in $[0, 20]$ as we have seen in the results of the experiments. A value higher than $10T$ indicates that user is more likely to be indoors while a value lower than $10T$ suggests that the user is outdoors.

The first idea that came to our mind was to set a linear function for the probabilities $p(y|Out) = \alpha y$. However this assumes that in the lower limit of the interval the device is for sure outdoors and in the upper limit it is certainly indoors.

This is not true because even outdoors the magnetic difference can be higher than $20T$ due to magnetic interferences and indoors the magnetic difference can be $0T$ in certain spots where the magnetic field measured is close to the Earth's magnetic field of that location. Therefore we use a corrected linear function $\beta + \alpha y$ for both the probability of being outdoors ($p(y|Out)$)and indoors ($p(y|In)$) for the magnetic field difference.

3.4.2.1 Magnetic Field Difference Probability of being Outdoors

The probability distribution of being outdoors for the magnetic field difference matches with the function $p(y|Out) = \beta + \alpha y$ and its integral from 0 to 20 should be one.

$$\int_0^{20} (\beta + \alpha y)dy = 1, \forall y \in [0, 20].$$

In this case we have two unknown variables α and β, so we have chosen them experimentally getting an α value of $\alpha = -0.00125$ and a β value of $\beta = 0.0625$. The resulting magnetic probability distribution of being outdoors is presented in Figure 29.

$$p(y|Out) = 0.0625 - 0.00125y, \forall y \in [0, 20].$$

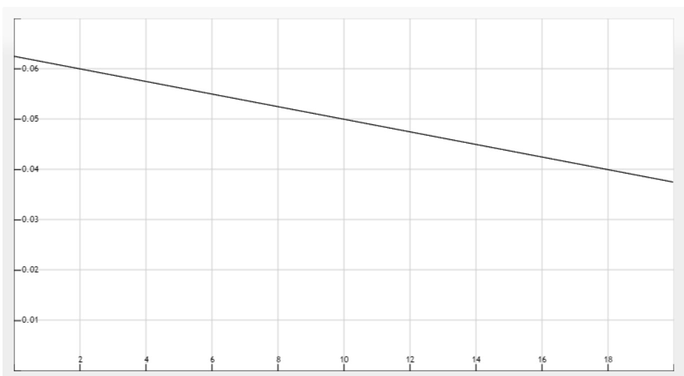

Figure 29: Magnetic field difference probability distribution of being outdoors.

3.4.2.2 Magnetic Field Difference Probability of being Indoors

Like the previous paragraph, we have to get the α and β parameters of the function $p(y|In) = \beta + \alpha y$ that satisfy the integral from 0 to 20 is one and represents the behaviour of the magnetic field difference indoors.

$$\int_0^{20} (\beta + \alpha y)dy = 1, \forall y \in [0, 20]$$

As before, we set the unknown variables experimentally getting an α value of $\alpha = 0.00125$ and a β value of $\beta = 0.0375$. This probability of being indoors is the opposing party of the previous distribution probability of being outdoors as it is shown in Figure 30.

$$p(y|In) = 0.0375 + 0.00125y, \forall y \in [0, 20].$$

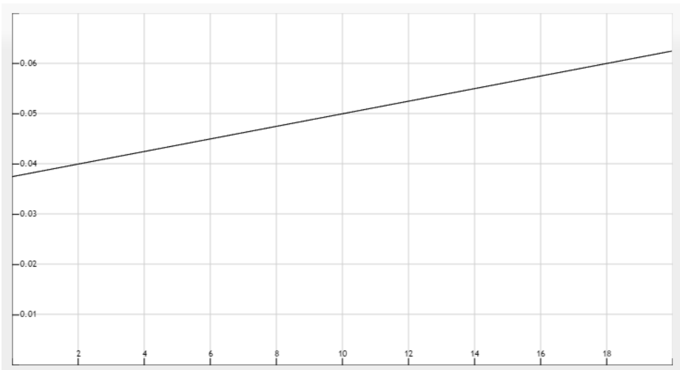

Figure 30: Magnetic field difference probability distribution of being indoors.

3.4.3 Number of Satellites Probability Distribution

The experiments with the GPS enabled (Section 3.3) indicates that with a number of fixed satellites close to 0 is more likely to be indoors while if it is close to 21 is more likely to be outdoors. We have considered that with less than 4 satellites you cannot be outdoors so we have set the interval in $[4, 20]$ for the probability distributions below.

3.4.3.1 Number of Satellites Probability of being Outdoors

If the number of satellites is less than or equal to 4, the probability distribution of being outdoors is 0 and it should rise with the number of the satellites. This increase is higher at the beginning due to is more likely to be outdoors with a number of satellites higher than 8.

For this reason we have chosen the square root $p(y|Out) = \alpha\sqrt{y-4}$ to represent the outdoor probability distribution and again we have to obtain α value that makes the integral one.

$$\int_4^{20} \alpha\sqrt{y-4}dy = 1, \forall y \in [4, 20]$$

The α value for this function represented in the Figure 31 is $\alpha = 0.0234$.

$$p(y|Out) = 0.0234\sqrt{y-4}, \forall y \in [4, 20].$$

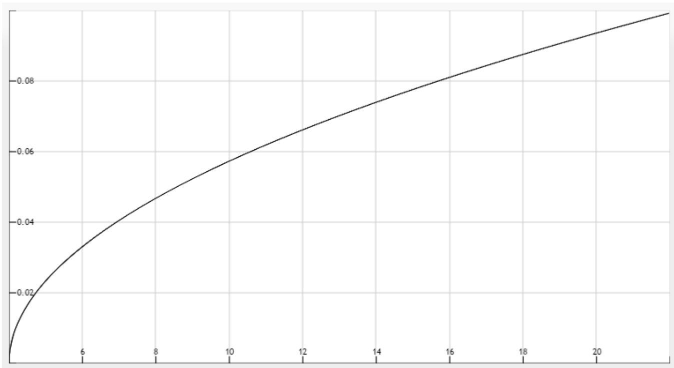

Figure 31: Number of satellites probability distribution of being outdoors.

3.4.3.2 Number of Satellites Probability of being Indoors

We have used the function $p(y|In) = \alpha(y-20)^2$ to match the probability distribution of being indoors with the behaviour of the number of satellites for the indoor-outdoor detection.

$$\int_4^{20} \alpha(y-20)^2 dy = 1, \forall y \in [4, 20]$$

The α value that satisfies the integral above is $\alpha = 0.000732$. As we can see in Figure 32, the number of satellites probability distribution of being indoors matches with the behaviour expected where a lower number of satellites implies higher probability of being indoors.

$$p(y|In) = 0.000732(y - 20)^2, \forall y \in [4, 20].$$

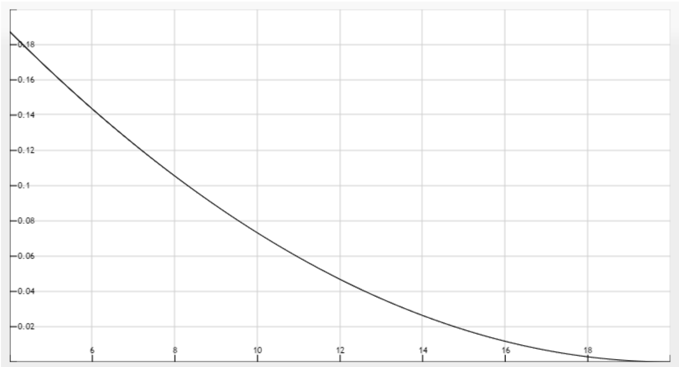

Figure 32: Number of satellites probability distribution of being indoors.

3.4.4 Signal to Noise Ratio Probability Distribution

This SNR is similar to the number of satellites and therefore we are going to use the same functions for the indoor and outdoor probability distributions but using the interval $[0, 36]$ obtained in the experiments.

3.4.4.1 Signal to Noise Ratio Probability of being Outdoors

The integral of the function $p(y|Out) = \alpha\sqrt{y}$ from $0dB$ to $36dB$ is one as the previous probability distributions.

$$\int_0^{36} \alpha\sqrt{y}dy = 1, \forall y \in [0, 36]$$

Solving this integral, we obtain an α value of $\alpha = 0.00694$ that results in the SNR probability distribution of being outdoors as shown in Figure 33 follows.

$$p(y|Out) = 0.00694\sqrt{y}, \forall y \in [0, 36].$$

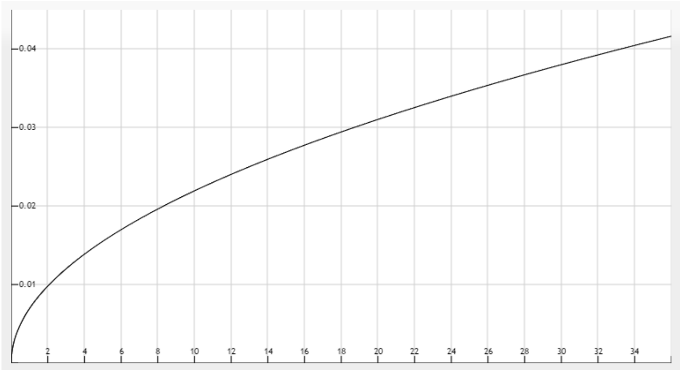

Figure 33: Signal to noise ratio probability distribution of being outdoors.

3.4.4.2 Signal to Noise Ratio Probability of being Indoors

The function used to express the SNR probability of being indoors is $p(y|In) = \alpha(y-36)^2$ meaning that the probability will decrease really fast from the higher value in $y = 0$ until $y = 36$ where the probability of being indoors is 0.

$$\int_0^{36} \alpha(y-36)^2 dy = 1, \forall y \in [0, 36].$$

The value $\alpha = 0.0000643$ satisfies the statement above, resulting a SNR probability distribution of being indoors like this (Figure 34):

$$p(y|In) = 0.0000643(y-36)^2, \forall y \in [0, 36].$$

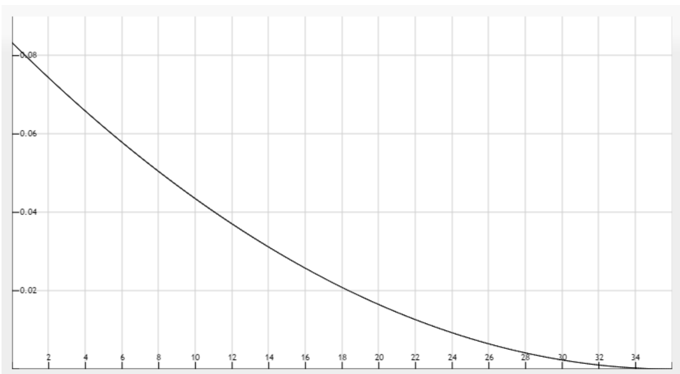

Figure 34: Signal to noise ratio probability distribution of being indoors.

3.4.5 Probability of being Outdoors

Once we have the conditional probabilities of each parameter we can write the joint probability function that contains all the necessary information of our Bayesian network,

$$P(Out, L, M, N_{Sat}, Snr) = P(Out)P(L|Out)P(M|Out)P(N_{Sat}|Out)P(Snr|Out).$$

To get the probability of being out based on the light intensity, the magnetic field difference, the number of satellites and the SNR, we proceed as in section 2.8 using the joint probability function of our model,

$$P(Out|L, M, N_{Sat}, Snr) = \frac{P(Out, L, M, N_{Sat}, Snr)}{P(Out, L, M, N_{Sat}, Snr) + P(\neg Out, L, M, N_{Sat}, Snr)},$$

where $P(\neg Out, L, M, N_{Sat}, Snr)$ is:

$$P(\neg Out)P(L|\neg Out)P(M|\neg Out)P(N_{Sat}|\neg Out)P(Snr|\neg Out).$$

We have set the probability of being out to $P(Out) = 0.5$ because we do not know a priori if we are outdoors or indoors. If $P(Out) = 0.5$, the probability of not being outdoors is $P(\neg Out) = 1 - 0.5 = 0.5$ that is the probability of being indoors $P(In)$.

$$P(In, L, M, N_{Sat}, Snr) = P(In)P(L|In)P(M|In)P(N_{Sat}|In)P(Snr|In)$$

Therefore the probability of being outdoors can be written as follows.

$$P(Out|L, M, N_{Sat}, Snr) = \frac{P(Out, L, M, N_{Sat}, Snr)}{P(Out, L, M, N_{Sat}, Snr) + P(In, L, M, N_{Sat}, Snr)}$$

We cannot forget that the conditional probabilities of each parameter depends on the value measured by the mobile phone. The table below summarizes the probability distributions of the model for each parameter.

Light Intensity	$P(L\|Out)$	$P(L\|In)$
$\forall y \in [0, 25]$	$0.012\sqrt{y}$	$0.000192(y-25)^2$

Magnetic Field Difference	$P(M\|Out)$	$P(M\|In)$
$\forall y \in [0, 20]$	$0.0625 - 0.00125y$	$0.0375 + 0.00125y$

Number of Satellites	$P(N_{Sat}\|Out)$	$P(N_{Sat}\|In)$
$\forall y \in [4, 20]$	$0.0234\sqrt{y-4}$	$0.000732(y-20)^2$

Signal to Noise Ratio	$P(Snr\|Out)$	$P(Snr\|In)$
$\forall y \in [0, 36]$	$0.00694\sqrt{y}$	$0.0000643(y-36)^2$

3.5 Algorithm for Indoor-Outdoor Detection

Although we know how to calculate the probability of being outdoors, this probability varies in each time-step due to the measurements of the device that change the conditional probabilities.

The values measured by the mobile phone concerning the light intensity, the magnetic field difference, the number of satellites and the SNR may change too much from one time-step to another. Typically is because of the state of user, that may change really fast from indoors to outdoors and vice versa. However some of this peaks appear due to a wrong measurement by the smartphone and if it is taken directly to calculate the outdoor probability, it may not match with the real state of the user.

To mitigate these peaks and get a more accurate outdoor probability, we have applied a low-pass filter to each measure obtained from the mobile phone in every time-step. Being x_n the current measurement and y_{n-1} the filtered value of the previous time-step, the filtered current measurement is y_n as follows:

$$y_n = \alpha x_n + (1-\alpha) y_{n-1}$$

Where α is a parameter between 0 and 1 that indicates how much the previous value influences the current value. We have set the α value in $\alpha = 0.5$ to reduce the peaks of the measurements but taking into account equally the current measurement and the previous filtered value.

With this correction, now we can obtain an accurate probability of being outdoors for each time-step as we explained in section 3.4. We need this outdoor probability to be as robust as possible so we decided to average the outdoor probabilities of five last time-steps ($P_i(Out|L, M, N_{Sat}, Snr) \forall i \in [1, 5]$) to get the final probability of being out ($P_{final}(Out|L, M, N_{Sat}, Snr)$).

$$P_{final}(Out|L, M, N_{Sat}, Snr) = \frac{\sum_{i=1}^{5} P_i(Out|L, M, N_{Sat}, Snr)}{5}$$

With all this we can now build our robust algorithm for indoor-outdoor detection which is a closed loop as we can see in Figure 35 where our output is the final probability of being out.

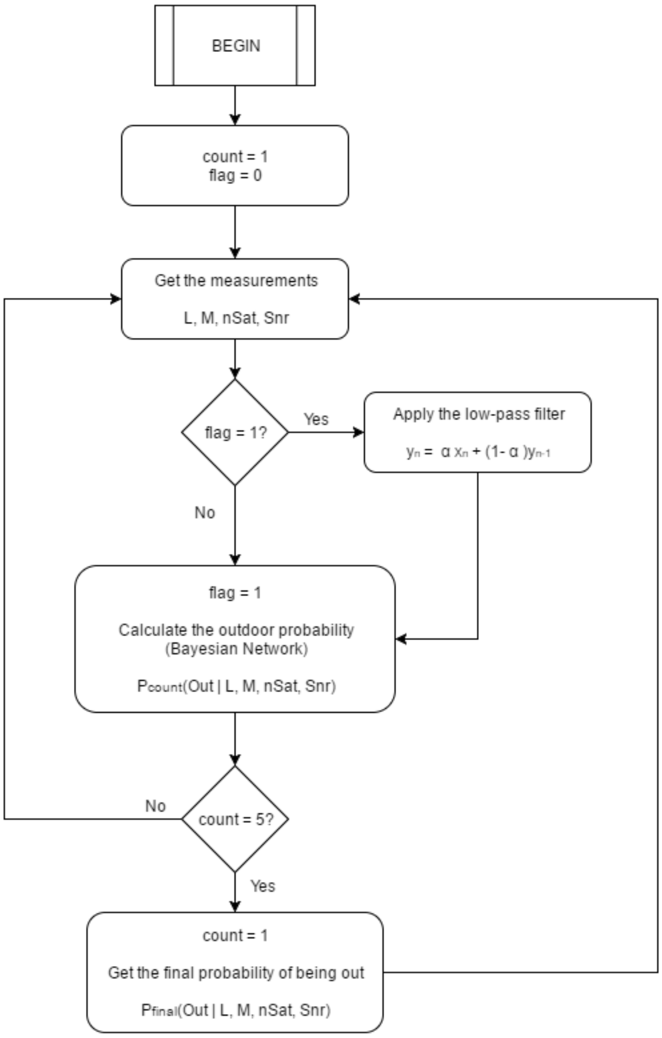

Figure 35: Flow chart of the indoor-outdoor detection algorithm.

We need two auxiliary variables for the algorithm: *count* and *flag*. The *count* variable is used to count the time-steps of the loop to get the final probability of being out based on five outdoor probabilities calculated with our Bayesian network. On the other hand the *flag* variable is used to apply the low-pass filter only when there is a previous value, in other words it allows the initialization of the system.

3.5.1 Test

Once we developed the algorithm for indoor-outdoor detection, we tested its functionality with the datasets recorded before.

The most complete dataset was recorded in Aalto University Campus (Espoo, Finland) on the 29-February-2016 at 11:00 am. It is a 5 minutes long recording where the user walked from the Main Building in Otakaari 1 to the Computer Science Building in Konemiehentie 2 (Figure 36).

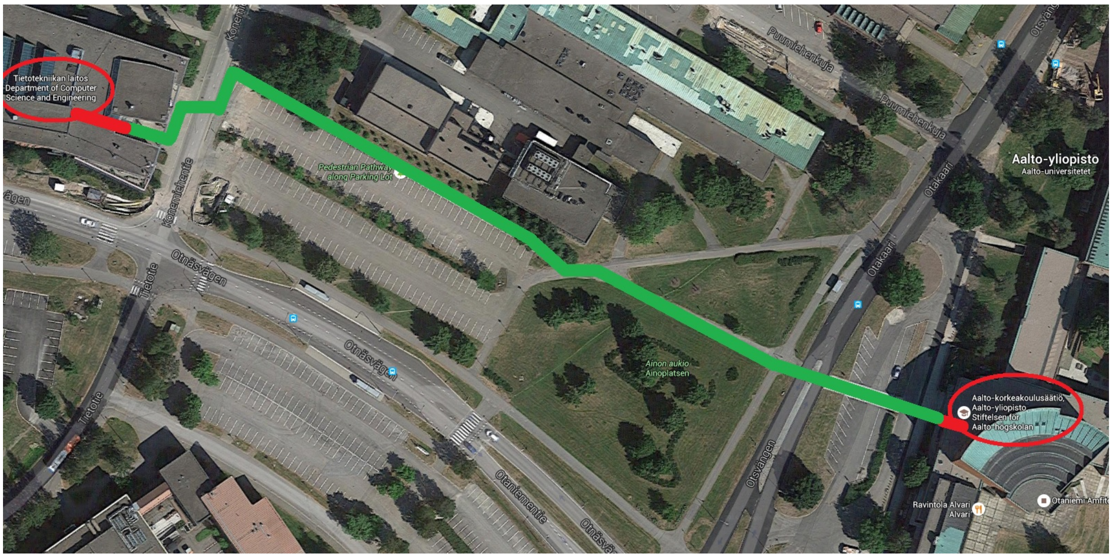

Figure 36: Route of the 29-February-2016 dataset. In red the parts of the rout indoors, in green the parts of the route outdoors. (Source: www.google.com/maps)

Figure 37 illustrates the probability of being out (vertical axis) along the time-steps (horizontal axis) gathered. The red spots represent the probability of being out in each time-step while the black stars show the average of 5 time-steps that correct the output of the algorithm.

The probability of being outdoors is close to 1 almost all the time the user is outdoors and close to 0 when the user is inside the buildings. Thus, we have set the threshold of being out in 0.5 meaning that if the final probability of being outdoors (corrected) is equal or higher than 0.5, our algorithm understands that the device is out and otherwise if it less than 0.5.

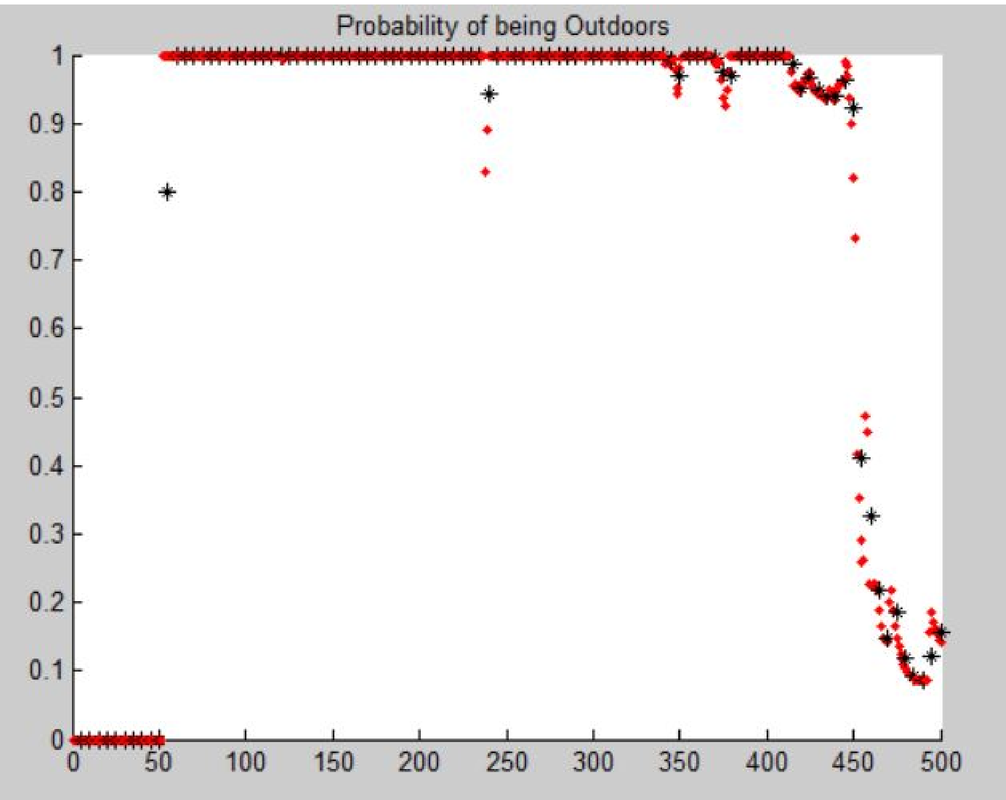

Figure 37: In red the probability of being outdoors in each time step, in black the final probability of being out taken every 5 time-steps.

The algorithm applied on the rest of the datasets presents a similar behaviour, with final probabilities of being out higher than 0.8 in outdoor environments and less than 0.3 indoors.

We noticed that during the transitions, the algorithm takes a while to detect the right state of the user (usually from 1 to 3 seconds). This is mainly due to the correction of the final probability that slows the response of the algorithm. However the average of the outdoors probabilities makes the algorithm more robust, fulfilling the aim of the research.

4 Implementation

As mentioned before, the final Android application has three main parts, the outdoor guidance, the indoor guidance and the indoor-outdoor transition. In this thesis we are going to explain the outdoor path planning and the transition between indoor and outdoor environments. In addition to do that we are going to describe how all these three parts are joined together and the functionality of the final Android application.

4.1 Scope

First of all we have to define the scope of the Android application to assure a right performance.

Figure 38: Huawei Honor 6. (Source: www.oppomart.com)

The final application was developed for the smartphone Huawei Honor 6 (Figure 38) that has the following specifications:

- Network technology: GSM/HSPA/LTE
- Platform: Android Operative System 5.1.1 (Lollipop).
- Memory:
  - Internal: 16 Gb and 3 Gb RAM (Random Access Memory).
  - External: 64 Gb microSD card.

- Communications:
  - WLAN
  - GPS
  - NFC
  - Bluetooth
- Sensors:
  - Ambient light sensor
  - Magnetometer
  - Proximity sensor
  - Accelerometer
  - Gyroscope

The application has been developed using the free software Android Studio [3] that provides all the necessary tools for building applications on Android devices. This software was installed on a personal laptop (Lenovo Y-50) that runs on Windows 10.

Furthermore, the indoor guidance (explained in another master thesis) has been developed with the IndoorAtlas SDK, that allows an accurate positioning inside buildings using magnetic field fingerprinting.

This technology uses the magnetic field measurements to make a geomagnetic map of a certain building without requiring any additional hardware besides a smarphone. Then it gets the indoor location of the user by comparing the magnetic field of the mobile phone against the geomagnetic map of the building.

This means that the indoor positioning inside a building can only be obtained once a geomagnetic map of the building is made. Thus, for our indoor-outdoor guidance application, we have to define the buildings that are already mapped where the indoor guidance can work.

Since the purpose of this application is to prove that it is possible to create an indoor-outdoor guidance, we have chosen only one building for the demo: the Sello mall in Leppävaarankatu 3-9 (Espoo, Finland).

[3]https://developer.android.com/studio/index.html

Figure 39: First floor of Sello mall.

Even though Sello has more than one floor, for the demo we are using only the first floor of the mall shown in Figure 39.

4.2 Outdoor Guidance

The most widespread location method for outdoor environments is the GPS as we have seen in Section 2.1.1 and therefore it is the technology we are going to use for building the outdoor guidance part of the Android application.

Our first idea was to use directly the Google Maps application, a free Android application (available in Google Play) that uses the GPS for outdoor positioning and routing. This application contains all the necessary resources to get the global position of the device and calculate any route to any goal chosen by the user.

However this application cannot be used from other Android application like ours, making impossible a smooth transition between the outdoor and indoor guidances. We used the Google Maps Direction API instead, because it contains the same resources as the Google Maps application that can be used from our application.

This part will explain how this Google Maps Direction API works to get the shortest path from your location to the goal of your choice. We are going to describe this process to make it understandable for anyone so we are not going to explain in detail the code of the application.

4.2.1 Google Maps Framework

The Google Maps framework is the first thing we need for the outdoor guidance because it provides the location of the device on the World's map. Thanks to Android Studio we can obtain this interface easily by creating a Google Maps class that consists basically on a map downloaded from the internet and several functions which are called when it is necessary. To use it, you need to register the application in Google Maps[4] to get a key that is saved in the manifest of the android application.

To get the location in real time (the blue dot in Figure 40) we have to set the *location request* to request a quality of service for location updates and then start this *location updates* when the class is called for the first time.

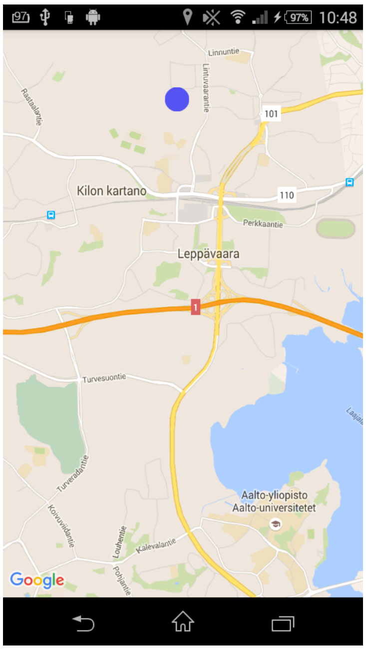

Figure 40: Google Maps Framework. The blue dot represents the location of the mobiles phone in real time.

Once we have our Google Maps class working we can proceed to get the route from the current position of the user to any location.

4.2.2 Google Maps Directions

This part is a little bit more complicated and needs the Google Maps Directions API. Once again you need to register the application to get the API key[5] (called *API_KEY*

[4] https://developers.google.com/maps/documentation/javascript/get-api-key
[5] https://developers.google.com/maps/documentation/directions/get-api-key

in the application) that is used to check the authenticity of the application.

The Google Maps Directions API is a service that calculates directions between locations using an HTTP request. You may specify the transportation mode to use such as driving, walking, bicycling and public transit. For the purpose of the application we are going to use the walking mode because the transition between indoors and outdoors will be done by foot.

The first step is to send the HTTP request when the Google Maps framework is called. As an example we are going to write the HTTP request of going from our current location (*Lat*, *Lng*) to Aalto University in Otakaari 1.

https://maps.googleapis.com/maps/api/directions/json?origin=Lat,Lng
&destination=Otakaari1&mode=walking&key=API_KEY;

The above example requests a JSON (JavaScript Object Notation) output. This file contains all the necessary information to walk from the current location to Aalto university. However we cannot get this information directly because this JSON file has a specific structure.

Therefore we have to parse this JSON file to paint the requested route. This is made thanks to a class called *DirectionsJSONParser* that returns the waypoints of the path. With all these waypoints, the last step is to paint the route in the Google Maps framework like it is shown in Figure 41.

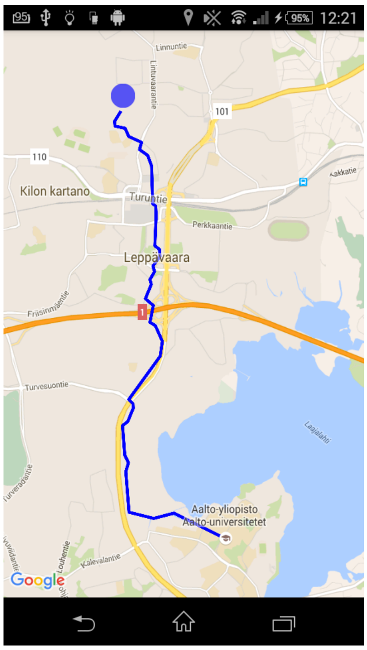

Figure 41: Google Maps Direction route from our current location (blue dot) to Aalto University.

4.3 Indoor-Outdoor Transition

This part is the core of the thesis because we have to apply the indoor-outdoor detection algorithm developed in Section 3.5 to a real time indoor-outdoor detection in an Android application.

Firstly we get the input variables of the algorithm by setting the following listeners on the Android application main class:

- **SensorEventListener**: This listener provides access to most of the sensors of the mobile phone such as the light sensor and the magnetometer. Thus we use the method called *OnSensorChanged* to get the values of the light intensity and the magnetic field every time they change.

  This method is called every time something happens in the measurements of the smartpohone that is pretty much every 1 millisecond so it is here where we call our algorithm method *main*. Nonetheless we are going to call the algorithm once every 0.5 seconds to avoid overloading the cell phone.

- **LocationListener**: This one allows the application to get the location of the device using the method *onLocationChanged*. This listener can be configured to use the network or the GPS provider. We have chosen the network provider because you do not need to enable the GPS of the mobile phone to get the estimated longitude and latitude of the user.

- **GpsStatus.Listener**: This listener has a method called *onGpsStatusChanged* that give us information about the GPS connexion status. We use it to count the number of satellites available every time something changes in the satellite status.

- **GpsStatus.NmeaListener**: As the previous one, this listener is used to get certain information about the GPS status but specifically for the NMEA data. It contains the method *onNmeaReceived* that give us once sentence at a time of the NMEA full message.

  After that we parse the sentences that start with *GPGSV* as we have explained in Section 3.3.3 to get the SNR average of all satellites available.

Once we have the inputs, we can implement the algorithm resulting the final probability of being outdoors as output. The objective of this real time detection is to be as robust as possible so we are going to consider all possible scenarios that may appear to apply the detection algorithm.

The algorithm works as described in Section 3.5 when all the parameters are available. But the algorithm does not take into account if some of these four parameters cannot be obtained.

For instance, the number of satellites and the SNR can only be obtained if the device has the GPS enabled. Notice that for the magnetic field difference we need the location of the user due to is an input for the magnetic field outdoors calculated with the dipole model of the Earth (see Section 3.3.1). However, we can obtain this location even if the GPS is not enabled because we use the network provider instead of the GPS provider for the positioning.

Furthermore the light intensity is only a relevant parameter when the transition occurs during the day, because during the night there is no sunlight and the probability distribution obtained cannot be applied.

To correct the final probability of being outdoors, we decided to consider an extra parameter for the Bayesian network that the literature study proves its reliability: the cell strength (Section 2.6).

4.3.1 Cell Strength

The cell towers are used to establish the communication between devices in phone calls. We can use the measurement of the cell strength to estimate if the user is indoors or outdoors. In our application we have used the listener **PhoneStateListener** to get this measurement by the method *onSignalStrengthsChanged*. In this way, every time the signal strength changes we obtain a new cell strength value.

The signal intensity can be measured in different manners, depending on the communication standard selected. We have chosen the GSM approach due to is the most common communication standard for cell phones. With the GSM standard, the signal strength obtained is a value between 0 and 31 dbm but for the probability distributions we decided to rescale the interval to [0,10] as we show in the following paragraphs.

4.3.1.1 Cell Strength Probability of being Outdoors

If the signal intensity is close to 0 means that the device cannot get the signal properly and therefore the probability of being outdoors is close to 0. On the other hand if the cell strength is close to 10, is more likely that the user is outdoors. These facts guide us to a linear probability distribution $p(y|Out) = \alpha y$ whose integral from 0 to 10 should be one,

$$\int_0^{10} \alpha y dy = 1, \forall y \in [0, 10].$$

The result is an α value of $\alpha = 0.2$ and therefore the linear probability distribution of being outdoors for the cell strength is (Figure 42):

$$p(y|Out) = 0.02y, \forall y \in [0, 10].$$

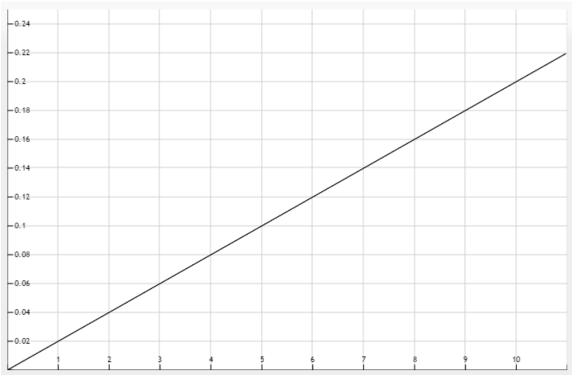

Figure 42: Cell strength probability distribution of being outdoors.

4.3.1.2 Cell Strength Probability of being Indoors

Indoors the probability distribution for the cell intensity is just the opposite of the probability distribution outdoors. It is next to 0 when the cell strength is around 10 and reaches the higher value when the cell strength is close to 0.

Hence the probability distribution of being indoors for the cell strength is (Figure 43):

$$p(y|In) = 0.2 - 0.02y, \forall y \in [0, 10].$$

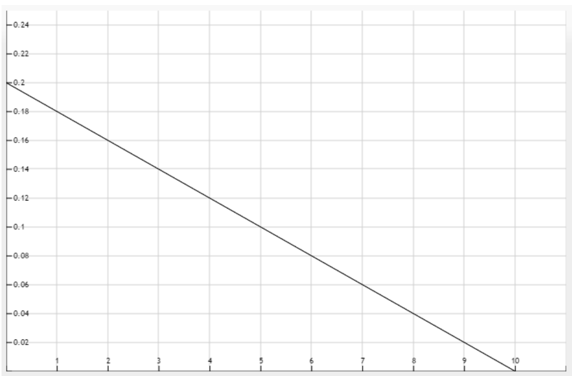

Figure 43: Cell strength probability distribution of being indoors.

4.3.2 Possible Scenarios

With the new parameter, the final Bayesian network has five components that are the magnetic field difference, the cell strength, the light intensity, the number of satellites and the signal to noise ratio as we can see in Figure 44.

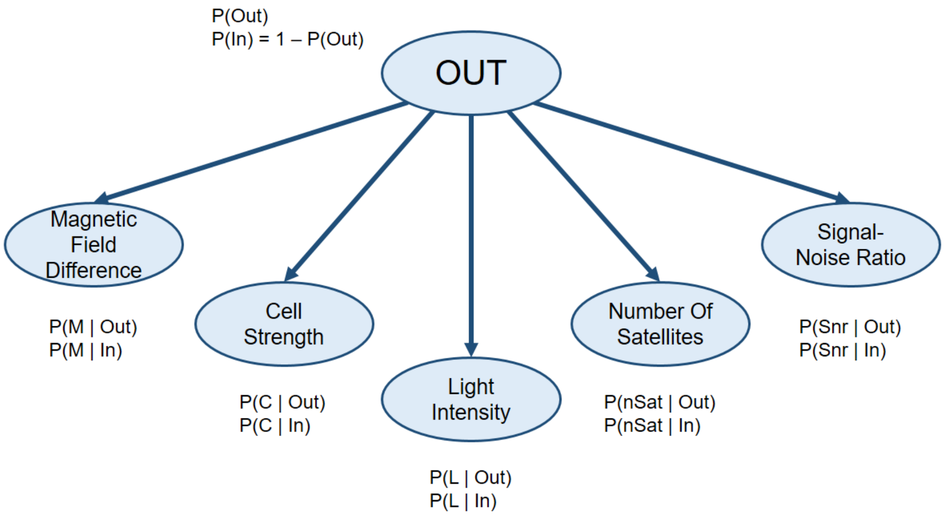

Figure 44: Final Bayesian network for indoor-outdoor detection.

Each component has its own probability distribution of being outdoors and indoors as we have described in previous sections. The next table summarizes the probability distributions for these five parameters.

Magnetic Field Difference	$P(M\|Out)$	$P(M\|In)$
$\forall y \in [0, 20]$	$0.0625 - 0.00125y$	$0.0375 + 0.00125y$

Cell Strength	$P(C\|Out)$	$P(C\|In)$
$\forall y \in [0, 10]$	$0.02y$	$0.2 - 0.02y$

Light Intensity	$P(L\|Out)$	$P(L\|In)$
$\forall y \in [0, 25]$	$0.012\sqrt{y}$	$0.000192(y - 25)^2$

Number of Satellites	$P(N_{Sat}\|Out)$	$P(N_{Sat}\|In)$
$\forall y \in [4, 20]$	$0.0234\sqrt{y-4}$	$0.000732(y - 20)^2$

Signal to Noise Ratio	$P(Snr\|Out)$	$P(Snr\|In)$
$\forall y \in [0, 36]$	$0.00694\sqrt{y}$	$0.0000643(y - 36)^2$

Consequently, being the joint probability function of being out:

$P(Out, M, C, L, N_{Sat}, Snr) =$

$$P(Out)P(M|Out)P(C|Out)P(L|Out)P(N_{Sat}|Out)P(Snr|Out)$$

And being the joint probability function of being indoors:

$P(In, M, C, L, N_{Sat}, Snr) =$

$$P(In)P(M|In)P(C|In)P(L|In)P(N_{Sat}|In)P(Snr|In)$$

With $P(Out) = P(In) = 0.5$, we can write the probability of being outdoors based on the relevant parameters mentioned before as:

$P(Out|M, C, L, N_{Sat}, Snr) =$

$$\frac{P(Out, M, C, L, N_{Sat}, Snr)}{P(Out, M, C, L, N_{Sat}, Snr) + P(In, M, C, L, N_{Sat}, Snr)}$$

This result can only be applied if we know the inputs of the Bayesian network, that are the light magnetic field, the cell strength, the light intensity, the number of satellites and the SNR.

The values that we can always obtain in real time are the magnetic field difference, the cell strength and the light intensity. The other two parameters, the number of satellites and the SNR can only be available if the GPS is enabled so for a robust indoor-outdoor detection we have to consider the possibility that the GPS is not enabled.

In addition, we notice the light intensity is only relevant during the daylight even though it is a parameter we can always get. Thus we have to consider that the user may use the application during the night, turning the light intensity an irrelevant component of the Bayesian network.

We can check easily if the GPS is enabled but a priori there is no measure that give us information about if the user is using the application during the day or during the night.

We decided to use the sunrise-sunset approach presented in Section 2.4.1 that gives us the sunrise and sunset times of the sun as result, based on the current day, month, year and latitude. Those values can be obtained from the mobile phone, including the latitude thanks to the network provider.

Once we compute the algorithm, we get two variables: the sunrise and sunset times. Check if the user is using the application during the day is to basically check if the current time of the device is between the sunrise and sunset times. Otherwise the user is using the application during the night.

This considerations lead us to three other possible scenarios apart from the general case presented before where we can obtain all the inputs.

- GPS enabled during the night:

  In this case, the light intensity is not relevant because during the night the light intensity outdoors has the same behaviour as indoors. Thus, the Bayesian network without considering the ambient light has the appearance of Figure 45.

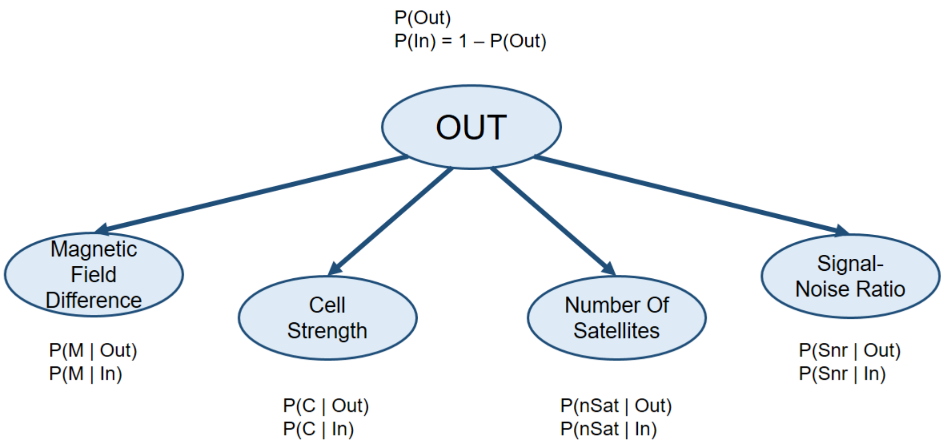

Figure 45: Bayesian network without taking into account the light intensity.

Then the probability of being outdoor based on the rest of the parameter is:
$P(Out|M, C, N_{Sat}, Snr) =$

$$\frac{P(Out, M, C, N_{Sat}, Snr)}{P(Out, M, C, N_{Sat}, Snr) + P(In, M, C, N_{Sat}, Snr)},$$

where $P(Out, M, C, N_{Sat}, Snr)$ is:

$$P(Out)P(M|Out)P(C|Out)P(N_{Sat}|Out)P(Snr|Out), \text{ and}$$

$P(In, M, C, N_{Sat}, Snr) =$ can be written as:

$$P(In)P(M|In)P(C|In)P(N_{Sat}|In)P(Snr|In).$$

- GPS not enabled during the day:

  Now we are going to specify the Bayesian network when the GPS is not enabled and we cannot obtain the number of satellites or the SNR parameters.

  If the detection is being done during the day, the parameters of the Bayesian network are the magnetic field difference, the cell strength and of course the light intensity (Figure 46).

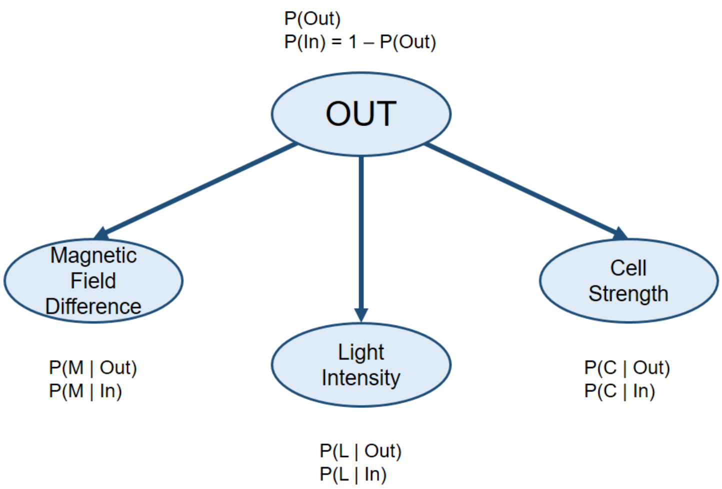

Figure 46: Bayesian network during the day without the GPS enabled.

Therefore the probability of being out is:

$$P(Out|M,C,L) = \frac{P(Out,M,C,L)}{P(Out,M,C,L) + P(In,M,C,L)},$$

where $P(Out,M,C,L)$ and $P(In,M,C,L)$ are respectively:

$$P(Out,M,C,L) = P(Out)P(M|Out)P(C|Out)P(L|Out),$$
$$P(In,M,C,L) = P(In)P(M|In)P(C|In)P(L|In).$$

- GPS not enabled during the night:

  In this last scenario, we are going to show what happens if the GPS is not enabled as the previous case, but during the night. Just the magnetic field difference and the cell strength are the important parameters because the rest of them are irrelevant, as we can observe in Figure 47.

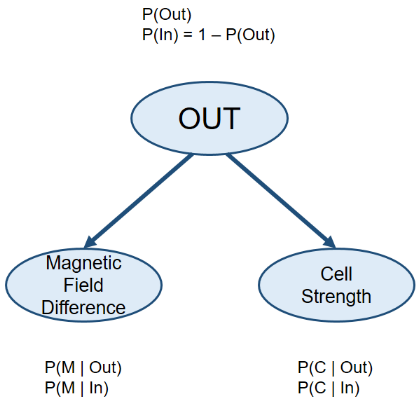

Figure 47: Bayesian network during the night without the GPS enabled.

As we can expect, the probability of being outdoors is really simple:

$$P(Out|M,C) = \frac{P(Out,M,C)}{P(Out,M,C) + P(In,M,C)}.$$

That can be written as:

$$P(Out|M,C) = \frac{P(Out)P(M|Out)P(C|Out)}{P(Out)P(M|Out)P(C|Out) + P(In)P(M|In)P(C|In)}.$$

4.3.3 Algorithm Performance

Following the algorithm presented in Section 3.5, we are going to assure the state of the user by calculating the outdoor probability average of five time-steps using the variable *count*.

First the algorithm waits until the network provider gives the current location of the device. Then it computes the sunrise-sunset algorithm to get the variable *Day* that has two possible values: true (1) if it is day and false (0) otherwise. After that it obtains the measurements from the Android listeners mentioned before and applies the low pass filter for each parameter.

At this point, it calculates an outdoor probability with the specific Bayesian network model depending on the availability of the GPS and the value of *Day*. The algorithm also uses the variable T to get a value for the outdoor probability every 0.5 seconds. When it has taken 5, the average of those 5 outdoor probabilities results the final output of the algorithm. The flow chart is shown in Figure 48.

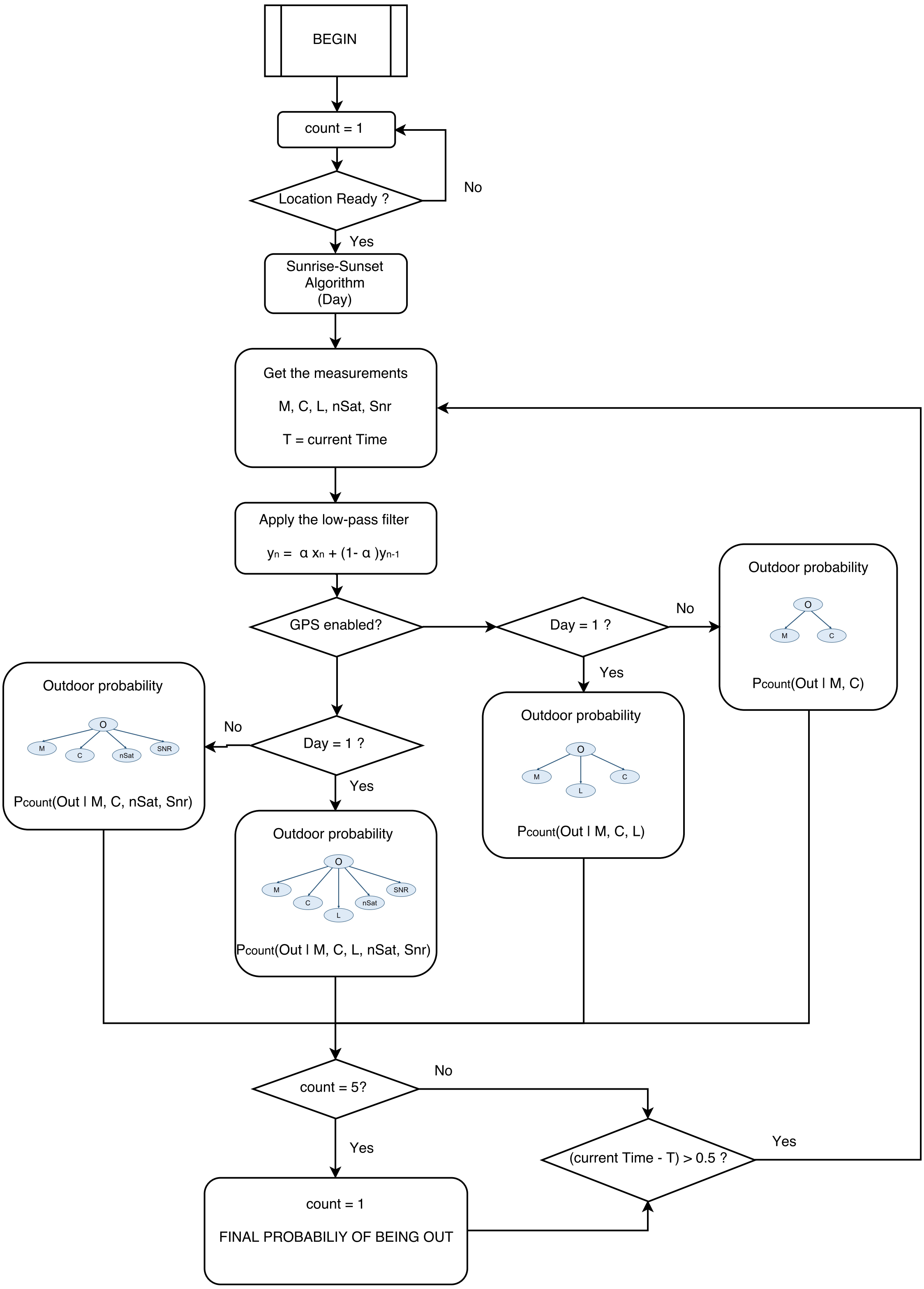

Figure 48: Flow of the implemented indoor-outdoor detection algorithm.

4.4 Final application

This section is going to explain the final Android application that is able to guide the user trough indoor and outdoor environments to reach the selected goal. As said, the application is divided in three parts, the outdoor guidance, the indoor-outdoor detection and the indoor guidance. We are going to explain how these parts work together to give the desire output.

The indoor-outdoor detection is always running in background because it is crucial to know when the transition between states occurs. Therefore we take for granted that the state of the user will be known at any time.

Concerning the indoor guidance, it can only proceed with the indoor location and routing if the building has been mapped previously. As mentioned, the application has been developed for the Sello mall so to assure the user is inside Sello, the device should be indoors (detection part) and the current location should be within the area of the mall. We have set a square area in Figure 49 that is large enough to guarantee we are inside Sello if we are indoors.

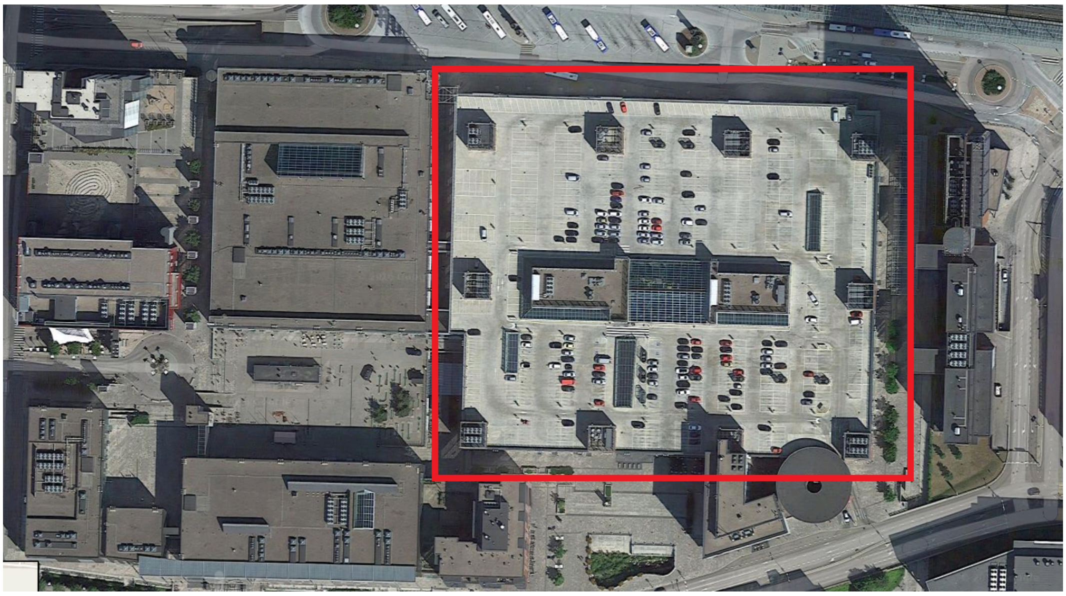

Figure 49: Area where you are inside Sello. (Source: www.google.com/maps)

When we run the application, we have three different initial states: outdoors, indoors outside Sello and inside Sello. For the guidance purpose, the outdoors and indoors outside Sello states work alike (Figure 50). You have two choices: go to a shop inside Sello or go somewhere else.

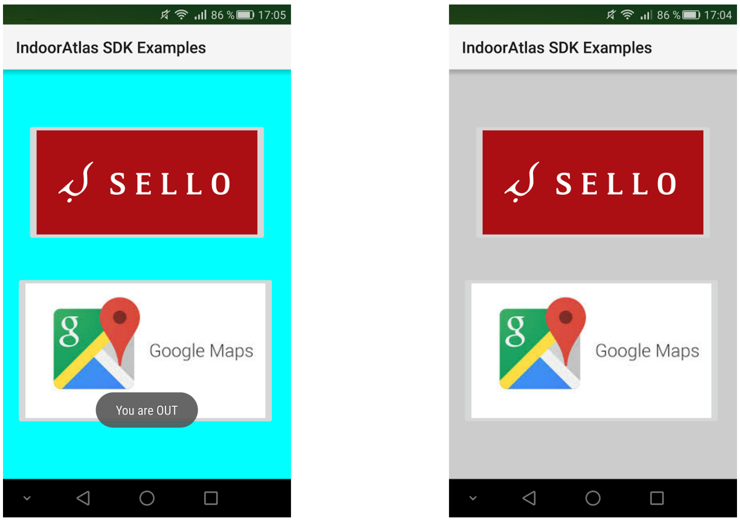

Figure 50: Left: screenshot of the application outdoors. Right: screenshot of the application indoors outside Sello.

If you select the Google Maps icon, the Google Maps application will be opened if it is installed on the device. Then the user may start the navigation by selecting the goal in the browser or by clicking in the map as it is shown in Figure 51.

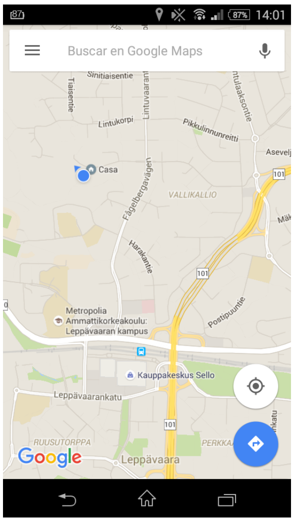

Figure 51: Screenshot of Google Maps application.

Google Maps uses the GPS to locate the device so if it is not enabled, the application will suggest to enable it. This application answers the problem of going from your current location outdoors to an outdoor location but cannot guide indoors as we have said previously.

On the other hand if you select the Sello icon, the guidance from outdoors to inside Sello will start as follows.

4.4.1 From Outdoors to Indoors

As an initial step in our guidance, the user has to select the goal inside Sello where he wants to go. This selection is really simple, you can either scroll down a shop in the list or search the name of the desire store in the browser. The Figure 52 illustrates an example where the user wants to go to the "Expert" store.

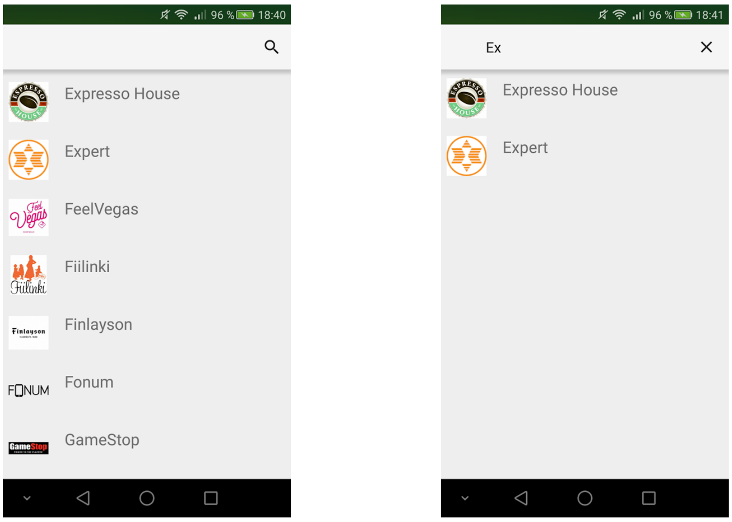

Figure 52: Left: screenshot of the list of goals. Right: screenshot of the filter.

When the goal is selected, a global variable is set to know in every step of the navigation where the user wants to go. Besides it start the outdoor navigation from the current location of the device to the mall, using the approach presented in Section 4.2 as we can observe in the Figure 53.

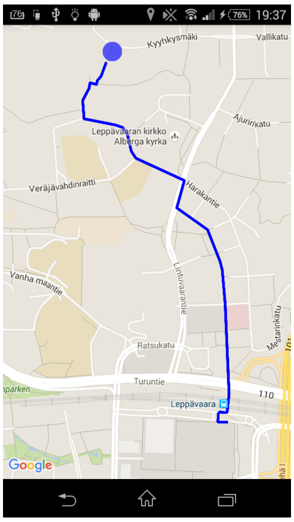

Figure 53: Screenshot of the outdoor guidance from the current location of the user (blue dot) to Sello.

We assume the user will follow the path drawn in the map to reach the entrance of the Sello mall. Afterwards, the application finds out the device is going indoors due to the indoor-outdoor detection algorithm that is working all the time in second plane.

However the user may be inside another building, so the application checks if the current location is within the area drawn before to know if the device is not only indoors but inside Sello.

If indeed the user has entered the mall, a pop-up is displayed for asking the user to switch to the indoor guidance (Figure 54). The user has two choices: select "No" to close the pop-up and continue the outdoor guidance or select "Yes" to open the indoor navigation from the current location to the store previous selected in the first step.

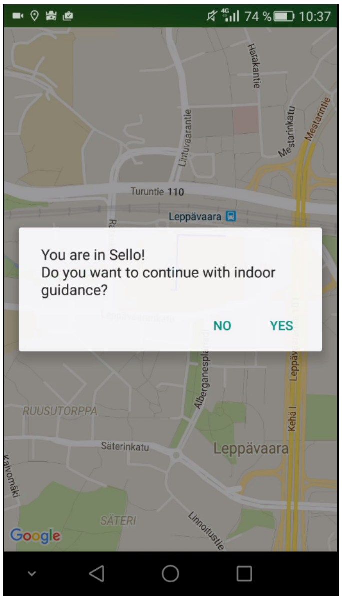

Figure 54: Screenshot of the switch pop-up from outdoors to indoors.

The indoor guidance starts with the positioning of the device indoors using the magnetic field. This process takes from 10 to 15 second, in this interval the user has to walk trough the corridors of the mall to get an accurate location inside Sello.

Once the accuracy of the location is high enough, the application computes an A star algorithm to draw the shortest path to the desire goal selected in the first step. An example of this performance can be found on Figure 55 where the user has entered the mall and wants to go to the "Expert" store.

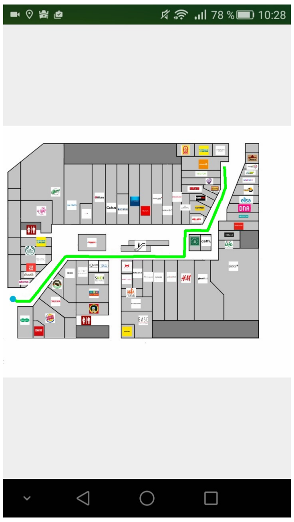

Figure 55: Screenshot of the indoor guidance from the current location of the user (blue dot) to the "Expert" store.

The path is updated if the user separates from the path, resulting a real time navigation through an indoor environment. In the next section we are going to analyse the guidance from inside Sello to outdoors.

4.4.2 From Indoors to Outdoors

When the initial state of the device is inside Sello, a different interface is displayed where you can choose between going to a shop inside Sello or writing a destination outdoors (Figure 56).

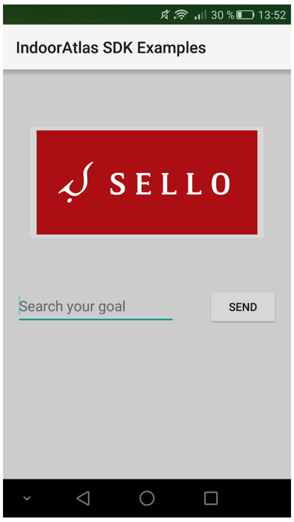

Figure 56: Screenshot of the application inside Sello.

If the user selects the Sello icon the list of shops available is shown as before. Once the goal is selected the indoor navigation screen is opened. Again the user has to walk for 10 to 15 second to get the accurate location inside Sello. When the accuracy has exceeded a certain value the path from that current position to the shop selected is displayed on the map of Sello as it is shown in Figure 57.

Figure 57: Screenshot of the indoors navigation inside Sello.

On the other hand, if the user wants to go somewhere else outside the mall, he can tip the goal in the text box and click the button "Send" to start the indoor-outdoor guidance.

When the location indoors is fixed, the next step is to go out the mall going to an exit gate. Thus, the application will draw the shortest path from the current location to the exit as we can observe in Figure 58.

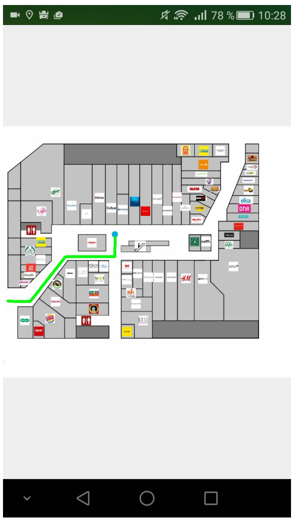

Figure 58: Screenshot of the indoors navigation to a store inside Sello.

Thanks to the detection part, when the user is leaving Sello the application notice the device is outdoors and asks the user if he wants to continue with the outdoor navigation (Figure 59).

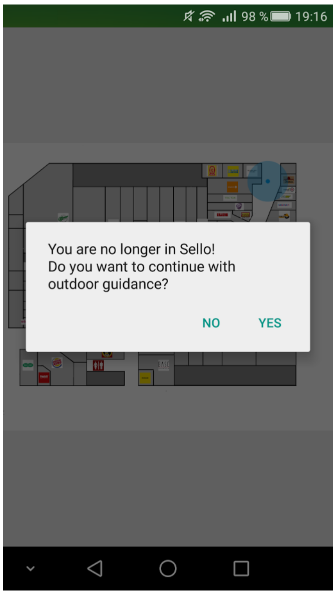

Figure 59: Screenshot of the switch pop-up from indoors to outdoors.

If the user clicks the "No" button the pop-up disappears and the application will try to locate the device indoors. Selecting the "Yes" button will open the outdoor navigation from the current location to the goal requested. For instance, the appearance of the navigation outdoors from Sello to Aalto Univeristy (Otakaari 1) is shown in Figure 60.

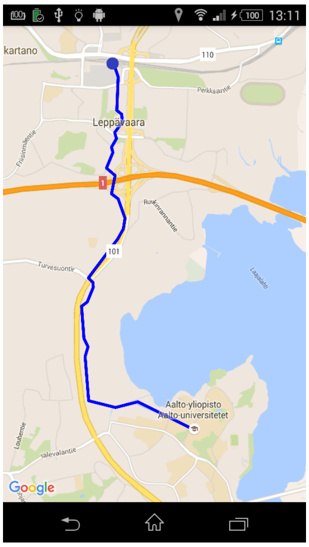

Figure 60: Screenshot of the outdoor navigation from Sello to Aalto University.

5 Results

It is difficult to evaluate the results of this thesis because the final outcome is the Android application for indoor-outdoor navigation itself.

Regarding the outdoor guidance, it has been developed using the Google Maps resources that has been confirmed as an effective tool for any GPS application. We are not going to take into account this outdoor navigation for analysis due to several studies has already proved the reliability of the Google Maps environment. Therefore this part will only analyse the performance of the application in the Sello mall, concerning the indoor-outdoor transition and vice versa.

First we are going to illustrate the differences between indoor and outdoor environments with Figure 61. On the top there is a picture of a place outdoors and on the bottom side a place indoors.

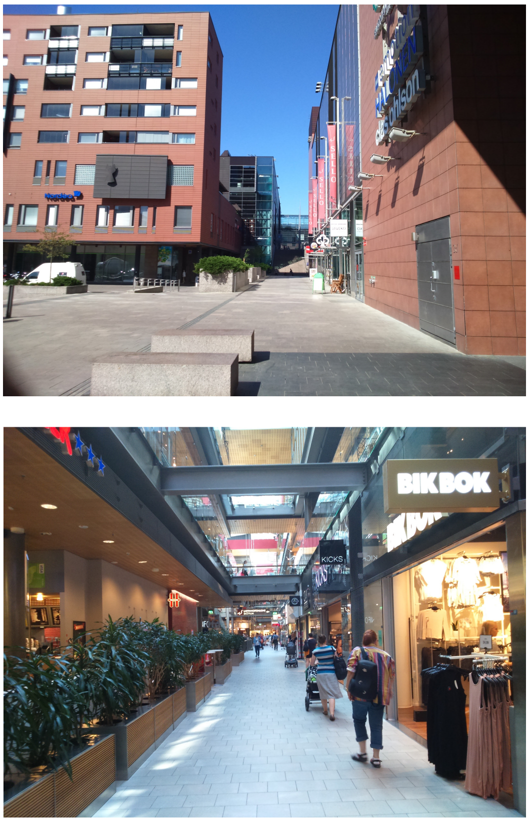

Figure 61: Top: outdoor environment. Bottom: indoor environment.

We can notice that in outdoor environments there is sunlight during the day, we can see the blue sky and there are not to many metal structures. On the other hand, inside buildings most of the light comes from artificial sources, we cannot see the sky due to the ceiling and there are plenty of metal components.

The analysis of the transitions have been done in two different gates of Sello. The gate number 1 is the west entrance of the building as we can see in the next picture.

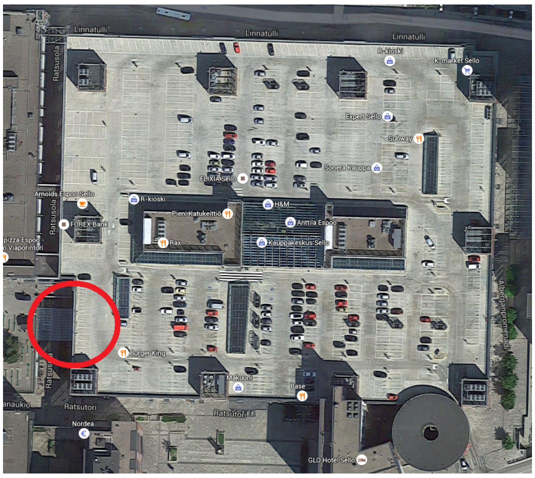

Figure 62: Caption of the west entrance of Sello.(Source: Google Maps)

It is used to access another building so it has a narrow outdoor corridor protected with a ceiling for the case it rains or snows. Indoors, the mall has an open area in front of the door and escalators to go up to the next floor (Figure 63).

Figure 63: Left: gate 1 outdoors. Right: gate 1 indoors.

This is the worst case we can find where the outdoor environment has components that can interfere with the detection such as ceiling, metal structures or shadow areas. However even in this scenario, the application detects the transitions between states thanks to the algorithm developed. To measure how well it detects the transitions, we decided to measure the elapsed time since the user changes the state until the application notices that change of state.

In this case, it takes from 3 to 5 seconds to detect the transition from indoors to outdoors and around 3 seconds from outdoors to indoors.

Moreover, the gate number 2 is the south entrance of the Sello mall.

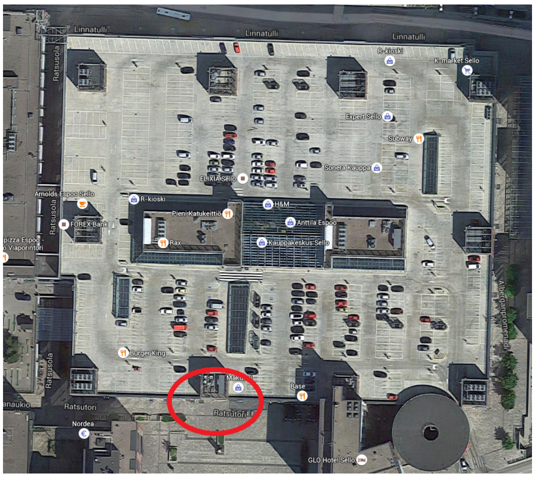

Figure 64: Caption of the south entrance of Sello.(Source: Google Maps)

In this case, Sello has a big square outdoors in front of gate 2 granting a free space without interferences. Besides, this gate lead us to a long narrow corridor indoors, plenty of artificial light sources (Figure 65).

Figure 65: Left: gate 2 outdoors. Right: gate 2 indoors.

This is indeed the best scenario, where outdoors there is no disturbance that can mislead the detection. The difference between indoor and outdoor states is really clear here, so the application achieves lower times to detect the change of state than the previous scenario.

From indoors to outdoors it takes around 2 seconds to detect the transition and from 1 to 3 seconds from outdoors to indoors.

6 Conclusions

We can conclude that we have successfully developed an Android application that allows the navigation trough indoor and outdoors environments. Since we have built only the outdoor guidance and the indoor-outdoor detection, we are not going into details with the indoor guidance.

The outdoor guidance is really accurate thanks to the resources we have used of Google Maps SDK. Furthermore we have tested that it is quite simple to get the directions from the current location of the user to any place desired by using the HTTP request via internet.

Regarding the indoor-outdoor detection we agreed that it is certainly robust due to the Bayesian network algorithm found in Section 3. The results of the previous section reveal that the detection from indoors to outdoors takes less than 4 seconds in 90% of cases and less than 3 seconds from outdoors to indoors in 100% of cases. This fact can be seen in Figure 66.

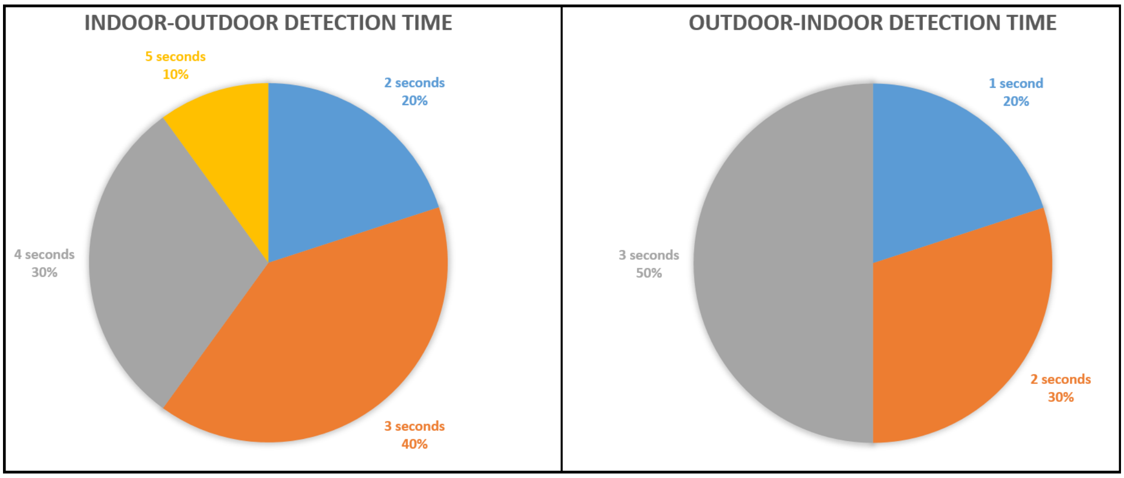

Figure 66: Left: time that it takes to detect the transition from indoors to outdoors. Right: time that it takes to detect the transition from outdoors to indoors.

7 Future lines of research

The application has been implemented only for the Sello mall and therefore a possible improvement of the system may be update our application with other buildings such as Aalto University or other public buildings. This line of research is more related with the indoor navigation but concerns the whole application.

In the current application the exit and entrance of Sello are fixed and even if you are south of the shopping center, the application will lead you to the west entrance set by default. The improvement here is to choose the entrance depending on where are you going and where do you came from.

Finally the indoor-outdoor detection using Bayesian networks can be improved in various aspects. The measurements of the Bayesian network parameters are directly related with the sensors and thus, the better sensors we have the more accurate is the value measured. Then more accurate sensors can improve the results of the algorithm developed, reducing the time of detection.

Other form of improvement is upgrading the probability distribution of the Bayesian network components to match better with the behaviour indoors and outdoors of each parameter.

References

Pontus Arfwedson and Joel Berglund. Indoor positioning system. Master's thesis, 2015.

Irad Ben-Gal. Bayesian networks (pdf). *Ruggeri, Fabrizio; Kennett, Ron S.; Faltin, Frederick W." Encyclopedia of Statistics in Quality and Reliability". Encyclopedia of Statistics in Quality and Reliability. John Wiley & Sons. doi*, 10:9780470061572, 2007.

Eugene Charniak. Bayesian networks without tears. Master's thesis, 1991.

Sudarshan S Chawathe. Beacon placement for indoor localization using bluetooth. Master's thesis, 2008.

Jaewoo Chung, Matt Donahoe, Chris Schmandt, Ig-Jae Kim, Pedram Razavai, and Micaela Wiseman. Indoor location sensing using geo-magnetism. Master's thesis, 2011.

Kevin Curran, Amanda Millar, and Conor Mc Garvey. Near field communication. *International Journal of Electrical and Computer Engineering*, 2(3):371, 2012.

Janne Haverinen and Anssi Kemppainen. Global indoor self-localization based on the ambient magnetic field. *Robotics and Autonomous Systems*, 57(10):1028–1035, 2009.

Sophie Hetet. Signal-to-noise ratio effects on the quality of gps observations. *Fredericton, New Brunswick: Department of Geodesy and Geomatics, University of New Brunswick*, 2000.

Richard Langley. Nmea 0183: A gps receiver. *GPS world*, 1995.

Jhong Sam Lee and Leonard E Miller. *CDMA systems engineering handbook*. Artech House, Inc., 1998.

Binghao Li, Ishrat J Quader, Andrew G Dempster, et al. On outdoor positioning with wi-fi. *Positioning*, 1(13), 2008.

Hui Liu, Houshang Darabi, Pat Banerjee, and Jing Liu. Survey of wireless indoor positioning techniques and systems. Master's thesis, 2007.

Rainer Mautz. Indoor positioning technologies. Master's thesis, Habilitationsschrift ETH Zürich, 2012, 2012.

Jean H Meeus. *Astronomical algorithms*. Willmann-Bell, Incorporated, 1991.

Esmond Mok and Günther Retscher. Location determination using wifi fingerprinting versus wifi trilateration. *Journal of Location Based Services*, 1(2):145–159, 2007.

Helmut Moritz. Geodetic reference system 1980. *Journal of Geodesy*, 54(3):395–405, 1980.

Michel Mouly, Marie-Bernadette Pautet, and Thomas Foreword By-Haug. *The GSM system for mobile communications*. Telecom publishing, 1992.

Judea Pearl and Stuart Russell. *Bayesian networks*. Computer Science Department, University of California, 1998.

Valentin Radu, Panagiota Katsikouli, Rik Sarkar, and Mahesh K Marina. A semi-supervised learning approach for robust indoor-outdoor detection with smartphones. Master's thesis, 2014.

Lenin Ravindranath, Calvin Newport, Hari Balakrishnan, and Samuel Madden. Improving wireless network performance using sensor hints. In *Proceedings of the 8th USENIX conference on Networked systems design and implementation*, pages 21–21. USENIX Association, 2011.

Anthony Rowe, Vikram Gupta, and Ragunathan Raj Rajkumar. Low-power clock synchronization using electromagnetic energy radiating from ac power lines. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*, pages 211–224. ACM, 2009.

Rob Van Dalen and Sergio Masferrer Oncala. Indoor/outdoor detection. Master's thesis, November 26 2013. US Patent 8,592,744.

Martin Walt. *Introduction to geomagnetically trapped radiation*. Cambridge University Press, 2005.

Ed Williams. Sunrire/sunset algorithm. http://williams.best.vwh.net/sunrise_sunset_algorithm.htm.

Pengfei Zhou, Yuanqing Zheng, Zhenjiang Li, Mo Li, and Guobin Shen. Iodetector: a generic service for indoor outdoor detection. Master's thesis, 2012.

Jean-Marie Zogg. Gps basics. Master's thesis, 2002.